STIC Search Report
Biotech-Chem Library

STIC Database Tracking Number: 177427

TO: Christian Fronda
Location: rem/2D78/2C70
Art Unit: 1652
Monday, January 30, 2006
Case Serial Number: 10/663033

From: Barb O'Bryen
Location: Biotech-Chem Library
Remsen 1a69
Phone: 571-272-2518

barbara.obryen@uspto.gov

Search Notes

Sequences created & searched:

Seq 3 fused to reverse complement of Seq 4
(3-revcomp-4)

Seq 4 reverse complement fused to Seq 3
(revcomp 4-seq 3)

gap size penalty changed from 0.5 (default)
to 0.1

STIC-Biotech/ChemLib

177427

From: Fronda, Christian
Sent: Tuesday, January 24, 2006 11:35 AM
To: STIC-Biotech/ChemLib
Subject: Sequence search for Serial No. 10/663,033

Importance: High

For Serial No. 10/663,033, there is a claim directed toward a polynucleotide obtained by PCR using PCR primers of SEQ ID NO: 3 and SEQ ID NO: 4.

Is there a way to search for polynucleotides that are obtained by PCR primers in the commercial, issued, PGPub, and pending databases? I only need results that have both SEQ ID NO: 3 and the **reverse complement of SEQ ID NO: 4.**

Please advise. Thank you.

Christian L. Fronda
Art Unit 1652
Office REM 2D78
Mailbox REM 2C70
(571)272-0929

****
Searcher:_____
Searcher Phone:_____
Date Searcher Picked up:_____
Date completed: 1/30
Searcher Prep Time:_____
Online Time:_____

****
Type of Search
NA#_____ AA#:_____
S/L:_____ Oligomer:_____
Encode/Transl:_____
Structure #:_____Text:_____
Inventor:_____ Litigation:___

****
Vendors and cost where applicable
STN:_____
DIALOG:_____
QUESTEL/ORBIT:_____
LEXIS/NEXIS:_____
SEQUENCE SYSTEM:_____
WWW/Internet:_____
Other (Specify):_____

OM nucleic - nucleic search, using sw model

Run on: January 29, 2006, 22:30:51 ; Search time 894 Seconds
(without alignments)
2479.748 Million cell updates/sec

Title: US-10-663-033-3-REVCOMP-4
Perfect score: 39
Sequence: 1 gaacactcacacgcgtcttatagcgagcgtctaattcta 39

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 0.1

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	24.2	62.1	325543	14	AC152657	AC152657 Bos tauru
c 2	22.6	57.9	208312	14	CR352328	CR352328 Danio rer
c 3	22	56.4	805	1	AF484440	AF484440 Helicobac
c 4	22	56.4	819	1	AF484441	AF484441 Helicobac
c 5	22	56.4	822	1	AF484437	AF484437 Helicobac
c 6	22	56.4	934	6	BD092531	BD092531 Identific
7	22	56.4	1037	1	AF093831	AF093831 Helicobac
8	22	56.4	1119	1	AF076779	AF076779 Helicobac
9	22	56.4	1119	6	AR154400	AR154400 Sequence
10	22	56.4	1119	6	BD238089	BD238089 Helicobac
11	22	56.4	1119	6	AR442784	AR442784 Sequence
12	22	56.4	1122	1	AF093828	AF093828 Helicobac
13	22	56.4	1124	1	AF093829	AF093829 Helicobac
14	22	56.4	1124	1	AF093832	AF093832 Helicobac
15	22	56.4	1124	1	AF093833	AF093833 Helicobac
16	22	56.4	1130	1	AF093830	AF093830 Helicobac
c 17	22	56.4	10530	1	AE000531	AE000531 Helicobac
18	22	56.4	166518	8	AC090514	AC090514 Homo sapi
c 19	21.6	55.4	3089	5	AJ719964	AJ719964 Gallus ga
20	21.6	55.4	102146	14	AC133396	AC133396 Felis cat
c 21	21.6	55.4	249683	14	AC119541	AC119541 Rattus no
c 22	21.4	54.9	110000	14	AP006493_4	Continuation (5 of
23	21.2	54.4	734	8	HSA333671	AJ333671 Homo sapi
24	21.2	54.4	101904	5	AL935310	AL935310 Zebrafish
c 25	21.2	54.4	152898	8	AL445240	AL445240 Human DNA
26	21.2	54.4	160737	14	AC026213	AC026213 Homo sapi
27	21.2	54.4	169675	14	AC146030	AC146030 Pan trogl
c 28	21.2	54.4	169856	14	AC144890	AC144890 Sus scrof
29	21.2	54.4	178009	14	AC119549	AC119549 Rattus no
c 30	21.2	54.4	220420	5	BX000452	BX000452 Zebrafish
c 31	21.2	54.4	318862	14	AC109547	AC109547 Rattus no
c 32	21	53.8	754	10	BV619750	BV619750 S216P6228
c 33	21	53.8	38688	14	AC100280	AC100280 Mus muscu
c 34	21	53.8	77087	8	AL355877	AL355877 Human DNA
c 35	21	53.8	80279	8	HS75K24	AL035700 Human DNA
36	21	53.8	150311	9	AC112081	AC112081 Rattus no
c 37	21	53.8	163500	14	AL450316	AL450316 Homo sapi
c 38	21	53.8	175172	14	AL355859	AL355859 Homo sapi
39	21	53.8	176135	9	AC161165	AC161165 Mus muscu
40	20.8	53.3	335	8	HS191XH12	Z23462 H. sapiens
c 41	20.8	53.3	95546	8	AC022816	AC022816 Homo sapi
c 42	20.6	52.8	373	10	BV094898	BV094898 RPAMMSEQ0
c 43	20.6	52.8	373	10	BV100363	BV100363 RPAMMSEQ0
c 44	20.6	52.8	373	10	BV156913	BV156913 RPAMMSEQ0
45	20.6	52.8	9033	7	BPAJ7641	AJ000741 Bacteriop

ALIGNMENTS

RESULT 1
AC152657
LOCUS AC152657 325543 bp DNA linear HTG 01-JUL-2005
DEFINITION Bos taurus clone CH240-10O13, *** SEQUENCING IN PROGRESS ***, 46 unordered pieces.
ACCESSION AC152657
VERSION AC152657.4 GI:68226917
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 325543)
AUTHORS Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D., Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,

Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,P., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.

TITLE Direct Submission

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 325543)

AUTHORS Worley,K.C.

TITLE Direct Submission

JOURNAL Submitted (16-NOV-2004) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE 3 (bases 1 to 325543)

AUTHORS Cow Genome Sequencing Consortium.

TITLE Direct Submission

JOURNAL Submitted (01-JUL-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT On Jun 26, 2005 this sequence version replaced gi:58037803.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

```
-------------- Genome Center
    Center: Baylor College of Medicine
    Center code: BCM
    Web site: http://www.hgsc.bcm.tmc.edu/
    Contact: hgsc-help@bcm.tmc.edu
-------------- Project Information
    Center project name: FAQI
    Center clone name: CH240-10O13
-------------- Summary Statistics
    Assembly program: Atlas 3.0;
    Consensus quality: 304082 bases at least Q40
    Consensus quality: 307999 bases at least Q30
    Consensus quality: 311731 bases at least Q20
    Estimated insert size: 312171; sum-of-contigs estimation
    Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
--------------
* NOTE: Estimated insert size may differ from sequence length
*   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)
* NOTE: This sequence may represent more than one clone.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 46 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*       1     1647: contig of 1647 bp in length
*    1648     1697: gap of 50 bp
*    1698     6641: contig of 4944 bp in length
*    6642     6691: gap of 50 bp
*    6692    19828: contig of 13137 bp in length
*   19829    19878: gap of 50 bp
*   19879    22273: contig of 2395 bp in length
*   22274    22373: gap of unknown length
*   22374    28672: contig of 6299 bp in length
*   28673    28722: gap of 50 bp
*   28723    30828: contig of 2106 bp in length
*   30829    30928: gap of unknown length
*   30929    33205: contig of 2277 bp in length
*   33206    33923: gap of 718 bp
*   33924    37464: contig of 3541 bp in length
*   37465    37564: gap of unknown length
*   37565    46565: contig of 9001 bp in length
*   46566    46615: gap of 50 bp
*   46616    54185: contig of 7570 bp in length
*   54186    54507: gap of 322 bp
*   54508    75814: contig of 21307 bp in length
*   75815    76332: gap of 518 bp
*   76333    77924: contig of 1592 bp in length
*   77925    78209: gap of 285 bp
*   78210    81979: contig of 3770 bp in length
*   81980    82029: gap of 50 bp
*   82030    83722: contig of 1693 bp in length
*   83723    83772: gap of 50 bp
*   83773    96928: contig of 13156 bp in length
*   96929    96978: gap of 50 bp
*   96979   105365: contig of 8387 bp in length
*  105366   105415: gap of 50 bp
*  105416   113285: contig of 7870 bp in length
*  113286   113517: gap of 232 bp
*  113518   134114: contig of 20597 bp in length
*  134115   135688: gap of 1574 bp
*  135689   142604: contig of 6916 bp in length
*  142605   142654: gap of 50 bp
*  142655   157716: contig of 15062 bp in length
*  157717   157766: gap of 50 bp
*  157767   166613: contig of 8847 bp in length
*  166614   166713: gap of unknown length
*  166714   171016: contig of 4303 bp in length
*  171017   171066: gap of 50 bp
*  171067   175916: contig of 4850 bp in length
*  175917   175966: gap of 50 bp
*  175967   179003: contig of 3037 bp in length
*  179004   179635: gap of 632 bp
*  179636   181102: contig of 1467 bp in length
*  181103   182232: gap of 1130 bp
*  182233   187896: contig of 5664 bp in length
*  187897   188286: gap of 390 bp
*  188287   197626: contig of 9340 bp in length
*  197627   197676: gap of 50 bp
*  197677   207375: contig of 9699 bp in length
*  207376   208252: gap of 877 bp
*  208253   215998: contig of 7746 bp in length
*  215999   216048: gap of 50 bp
*  216049   236221: contig of 20173 bp in length
*  236222   236271: gap of 50 bp
*  236272   254150: contig of 17879 bp in length
*  254151   254200: gap of 50 bp
*  254201   261491: contig of 7291 bp in length
*  261492   261541: gap of 50 bp
*  261542   275745: contig of 14204 bp in length
*  275746   275845: gap of unknown length
*  275846   279904: contig of 4059 bp in length
*  279905   279954: gap of 50 bp
*  279955   283386: contig of 3432 bp in length
*  283387   283436: gap of 50 bp
```

```
            *   283437   292331: contig of 8895 bp in length
            *   292332   292381: gap of 50 bp
            *   292382   295701: contig of 3320 bp in length
            *   295702   296854: gap of 1153 bp
            *   296855   305830: contig of 8976 bp in length
            *   305831   306127: gap of 297 bp
            *   306128   310233: contig of 4106 bp in length
            *   310234   310285: gap of 52 bp
            *   310286   317330: contig of 7045 bp in length
            *   317331   317430: gap of unknown length
            *   317431   318430: contig of 1000 bp in length
            *   318431   318530: gap of unknown length
            *   318531   319535: contig of 1005 bp in length
            *   319536   319635: gap of unknown length
            *   319636   320712: contig of 1077 bp in length

  Query Match             62.1%;  Score 24.2;  DB 14;  Length 325543;
  Best Local Similarity   78.4%;  Pred. No. 3.8;
  Matches   29;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          3 ACACTCACACGCGTCTTATAGCGAGCGTCTAATTCTA 39
              | ||||| || | |||| |||| | |||||||||| ||
Db      18261 AGACTCATACACATCTTCTAGCTATCGTCTAATTTTA 18297


RESULT 2
CR352328/c
LOCUS       CR352328              208312 bp    DNA     linear   HTG 05-MAR-2005
DEFINITION  Danio rerio clone CH211-212K18, WORKING DRAFT SEQUENCE, 3 unordered
            pieces.
ACCESSION   CR352328
VERSION     CR352328.5  GI:60543659
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 208312)
  AUTHORS   McLaren,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-MAR-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Mar 5, 2005 this sequence version replaced gi:46981214.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            -------------- Project Information
            Center project name: zC212K18
            -------------- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Chemistry: Dye-terminator; 100% of reads
            Consensus quality: 207646 bases at least Q40
            Consensus quality: 207804 bases at least Q30
            Consensus quality: 207867 bases at least Q20
            Insert size: 208112; sum-of-contigs
            Insert size: 206965; 7.7% error; agarose-fp
            Quality coverage: 10.35x in Q20 bases; sum-of-contigs Quality
            coverage: 10.53x in Q20 bases; agarose-fp
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 3 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1   101194: contig of 101194 bp in length
            *   101195   101294: gap of 100 bp
            *   101295   192622: contig of 91328 bp in length
            *   192623   192722: gap of 100 bp
            *   192723   208312: contig of 15590 bp in length.
FEATURES             Location/Qualifiers
     source          1. .208312
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="CH211-212K18"
                     /clone_lib="CHORI-211"
     misc_feature    1. .101194
                     /note="assembly_fragment:01920
                     fragment_chain:1
                     clone_end:SP6
                     vector_side:left"
     misc_feature    101295. .192622
                     /note="assembly_fragment:00203
                     fragment_chain:1
                     clone_end:T7
                     vector_side:right"
     misc_feature    192723. .208312
                     /note="assembly_fragment:00017"
ORIGIN

  Query Match             57.9%;  Score 22.6;  DB 14;  Length 208312;
  Best Local Similarity   75.7%;  Pred. No. 22;
  Matches   28;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 GAACACTCACACGCGTCTTATAGCGAGCGTCTAATTC 37
              |||||||| || |  || || |||||  ||||| |||
Db     200088 GAACACTCCCAGGTTTCCTACAGCGAAGGTCTACTTC 200052


RESULT 3
AF484440/c
LOCUS       AF484440                805 bp    DNA     linear   BCT 05-MAY-2005
DEFINITION  Helicobacter pylori isolate 60190 HpyIIIM protein (hpyIIIM) gene,
            partial cds.
ACCESSION   AF484440
VERSION     AF484440.1  GI:33337332
KEYWORDS    .
SOURCE      Helicobacter pylori
  ORGANISM  Helicobacter pylori
            Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
            Helicobacteraceae; Helicobacter.
REFERENCE   1  (bases 1 to 805)
  AUTHORS   Ando,T., Blaser,M.J. and Wassenaar,T.M.
  TITLE     Evolutionary history of hrgA, a gene replacing the restriction gene
            in an R-M gene locus of H. pylori
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 805)
  AUTHORS   Ando,T., Blaser,M.J. and Wassenaar,T.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-FEB-2002) First Department of Internal Medicine,
            Nagoya University School of Medicine, 65 Tsurumai-cho Showa-ku,
            Nagoya, Aichi 466-8550, Japan
FEATURES             Location/Qualifiers
     source          1. .805
                     /organism="Helicobacter pylori"
                     /mol_type="genomic DNA"
                     /isolate="60190"
                     /db_xref="taxon:210"
     gene            1. .>805
                     /gene="hpyIIIM"
     CDS             1. .>805
                     /gene="hpyIIIM"
                     /codon_start=1
                     /transl_table=11
                     /product="HpyIIIM protein"
                     /protein_id="AAQ13334.1"
                     /db_xref="GI:33337333"
                     /translation="MKASCLSRPIFKSQDQLFTLYQGDCNEVLPQFENAFDLIFADPP
                     YFLSNDGLSIQSGKIVSVNKGDWDKENGINDIDEFNCQWINNAKKALKNTGSLLISGT
```

YHNIFSLGRILQKLDFKILNLITWQKTNPPPNFSCRYLTHSAEQIIWARKSRKHKHVF
NYEVLKKINNDKQMRDVWSFPAIAPWEKANGKHPTQKPLALLVRLLLMASDGNSLIGD
PFSGSSTTGIAANLLKRQFIGIEKESEFIKISMNRKLELDARYKKIRSKI"

ORIGIN

Query Match 56.4%; Score 22; DB 1; Length 805;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 TTATAGCGAGCGTCTAATTCTA 39
||||||||||||||||||||||
Db 785 TTATAGCGAGCGTCTAATTCTA 764

RESULT 4
AF484441/c
LOCUS AF484441 819 bp DNA linear BCT 05-MAY-2005
DEFINITION Helicobacter pylori isolate 9627 HpyIIIM protein (hpyIIIM) gene, partial cds.
ACCESSION AF484441
VERSION AF484441.1 GI:33337334
KEYWORDS .
SOURCE Helicobacter pylori
ORGANISM Helicobacter pylori
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales; Helicobacteraceae; Helicobacter.
REFERENCE 1 (bases 1 to 819)
AUTHORS Ando,T., Blaser,M.J. and Wassenaar,T.M.
TITLE Evolutionary history of hrgA, a gene replacing the restriction gene in an R-M gene locus of H. pylori
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 819)
AUTHORS Ando,T., Blaser,M.J. and Wassenaar,T.M.
TITLE Direct Submission
JOURNAL Submitted (19-FEB-2002) First Department of Internal Medicine, Nagoya University School of Medicine, 65 Tsurumai-cho Showa-ku, Nagoya, Aichi 466-8550, Japan
FEATURES Location/Qualifiers
source 1. .819
/organism="Helicobacter pylori"
/mol_type="genomic DNA"
/isolate="9627"
/db_xref="taxon:210"
gene 1. .>819
/gene="hpyIIIM"
CDS 1. .>819
/gene="hpyIIIM"
/codon_start=1
/transl_table=11
/product="HpyIIIM protein"
/protein_id="AAQ13335.1"
/db_xref="GI:33337335"
/translation="MKTPCLSRPIFKSQDQLFTLYQGDCNEVLPQFENRFDLIFADPP
YFLSNDGLSIQSGKIVSVNKGDWDKEEGVNGIDEFNYQWINNAKKALKNTGSLLISGT
YHNIFSLGRLLQKLDFKILNLITWQKTNPPPNFSCRYLTHSAEQIIWARKSYKHKHVF
NYEVLKKINNDKQMRDVWSFPAIAPWEKVNGKHPTQKPLALLVRLLLMASDENSLIGD
PFSWSSTTGIAANLLKREFIGIEKESGFIKISMDRKIELDARYKEIRSKIKDLNH"

ORIGIN

Query Match 56.4%; Score 22; DB 1; Length 819;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 TTATAGCGAGCGTCTAATTCTA 39
||||||||||||||||||||||
Db 785 TTATAGCGAGCGTCTAATTCTA 764

RESULT 5
AF484437/c
LOCUS AF484437 822 bp DNA linear BCT 05-MAY-2005
DEFINITION Helicobacter pylori isolate JP2 HpyIIIM protein (hpyIIIM) gene, partial cds.
ACCESSION AF484437
VERSION AF484437.1 GI:33337326
KEYWORDS .
SOURCE Helicobacter pylori
ORGANISM Helicobacter pylori
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales; Helicobacteraceae; Helicobacter.
REFERENCE 1 (bases 1 to 822)
AUTHORS Ando,T., Blaser,M.J. and Wassenaar,T.M.
TITLE Evolutionary history of hrgA, a gene replacing the restriction gene in an R-M gene locus of H. pylori
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 822)
AUTHORS Ando,T., Blaser,M.J. and Wassenaar,T.M.
TITLE Direct Submission
JOURNAL Submitted (19-FEB-2002) First Department of Internal Medicine, Nagoya University School of Medicine, 65 Tsurumai-cho Showa-ku, Nagoya, Aichi 466-8550, Japan
FEATURES Location/Qualifiers
source 1. .822
/organism="Helicobacter pylori"
/mol_type="genomic DNA"
/isolate="JP2"
/db_xref="taxon:210"
gene 1. .>822
/gene="hpyIIIM"
CDS 1. .>822
/gene="hpyIIIM"
/codon_start=1
/transl_table=11
/product="HpyIIIM protein"
/protein_id="AAQ13331.1"
/db_xref="GI:33337327"
/translation="MEVPYLSRPIFKSQDQLFTLYQGDCNEVLPQFENQFDLIFADPP
YFLSNDGLSIQSGKIVSVNKGDWDKEESVNDIDEFNYQWINNAKKALKNTGSLLISGT
YHNLFSLGRILQKLDFKILNLITWQKTNPPPNFSCRYLTHSAEQIIWARKSHKHKHVF
NYEVLKKINNDKQMRDVWSFPAIAPWEKVNGKHPTQKPLALLVRLLLMASDENSLIGD
PFSGSSTTGIAANLLKRQFIGIEKESEFIQISMDRKLELDARYKEIRSKIKDLNLK"

ORIGIN

Query Match 56.4%; Score 22; DB 1; Length 822;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 TTATAGCGAGCGTCTAATTCTA 39
||||||||||||||||||||||
Db 785 TTATAGCGAGCGTCTAATTCTA 764

RESULT 6
BD092531/c
LOCUS BD092531 934 bp DNA linear PAT 27-AUG-2002
DEFINITION Identification of polynucleotides encoding novel helicobacter polypeptides in the helicobacter genome.
ACCESSION BD092531
VERSION BD092531.1 GI:22638142
KEYWORDS JP 2001527393-A/312.
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 934)
AUTHORS Kleanthous,H., Garawi,A.A., Miller,C., Tomb,J.F. and Oomen,R.P.
TITLE Identification of polynucleotides encoding novel helicobacter polypeptides in the helicobacter genome
JOURNAL Patent: JP 2001527393-A 312 25-DEC-2001;
MERIEUX ORAVAX SOCIETE EN NOM COLLECTIF PASTEUR MERIEUX SERUMS ET VACCINS AGROBIOLOGICAL RESOURCES MINISTRY O SA, HUMAN GENOME SCIENCES INC
COMMENT PN JP 2001527393-A/312
PD 25-DEC-2001
PF 01-APR-1998 JP 1998541947
PR 01-APR-1997 US 08/833457,24-JUN-1997 US 08/881227 PR 29-JUL-1997 US 08/902615

```
            PI   HAROLD KLEANTHOUS,AMAL AL GARAWI,CHARLES MILLER,JEAN FRANCOIS
            PI   TOMB,
            PI   RAYMOND PETER OOMEN
            PC   A01N43/04,A61K31/70
            CC   Strandedness: Single;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers.
FEATURES             Location/Qualifiers
     source          1. .934
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match           56.4%;  Score 22;  DB 6;  Length 934;
  Best Local Similarity  100.0%;  Pred. No. 53;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         18 TTATAGCGAGCGTCTAATTCTA 39
              ||||||||||||||||||||||
Db        844 TTATAGCGAGCGTCTAATTCTA 823


RESULT 7
AF093831
LOCUS       AF093831               1037 bp    DNA     linear   BCT 13-MAY-1999
DEFINITION  Helicobacter pylori strain UA1218 alpha-1,2-fucosyltransferase
            gene, complete cds.
ACCESSION   AF093831
VERSION     AF093831.1  GI:4808594
KEYWORDS    .
SOURCE      Helicobacter pylori
  ORGANISM  Helicobacter pylori
            Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
            Helicobacteraceae; Helicobacter.
REFERENCE   1  (bases 1 to 1037)
  AUTHORS   Wang,G., Rasko,D.A., Sherburne,R. and Taylor,D.E.
  TITLE     Molecular genetic basis for the variable expression of Lewis Y
            antigen in Helicobacter pylori: analysis of the alpha (1,2)
            fucosyltransferase gene
  JOURNAL   Mol. Microbiol. 31 (4), 1265-1274 (1999)
   PUBMED   10096092
REFERENCE   2  (bases 1 to 1037)
  AUTHORS   Wang,G. and Taylor,D.E.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-SEP-1998) Medical Microbiology and Immunology,
            University of Alberta, 1-51 Medical Science Building, Edmonton, AB
            T6G 2H7, Canada
FEATURES             Location/Qualifiers
     source          1. .1037
                     /organism="Helicobacter pylori"
                     /mol_type="genomic DNA"
                     /strain="UA1218"
                     /db_xref="taxon:210"
     CDS             58. .957
                     /codon_start=1
                     /transl_table=11
                     /product="alpha-1,2-fucosyltransferase"
                     /protein_id="AAD29867.1"
                     /db_xref="GI:4808595"
                     /translation="MAFKVVQICGGLGNQMFQYAFAKSLQKHSNTPVLLDTTSFDWSN
                     RKMQLELFPIDLPYASEKEIAIAKMQHLPKLVREALKYIGFDRVSQEIVFEYEPKLLK
                     PSRLTYFFGYFQDPRYFDAISPLIKKTLTLPPPPRKGNNKKKEEEYHRKLELILAAKN
                     GVFVHIRRGDYVGIGCQLGIDYQKKALEYMAKRVPNMELFVFCEDLEFTQNLDLGYPF
                     TDMTTRDKDEEAYWDMLLMQSCQHGIIANSTYSWWAAYLINNPEKIIIGPKHWLFGHE
                     NILCKEWVKIESHFEVKSQKYNA"
ORIGIN

  Query Match           56.4%;  Score 22;  DB 1;  Length 1037;
  Best Local Similarity  100.0%;  Pred. No. 53;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         18 TTATAGCGAGCGTCTAATTCTA 39
              ||||||||||||||||||||||
Db       1016 TTATAGCGAGCGTCTAATTCTA 1037


RESULT 8
AF076779
LOCUS       AF076779               1119 bp    DNA     linear   BCT 12-MAR-1999
DEFINITION  Helicobacter pylori alpha-1,2-fucosyltransferase (fucT2) gene,
            fucT2-type I allele, complete cds.
ACCESSION   AF076779
VERSION     AF076779.1  GI:4093138
KEYWORDS    .
SOURCE      Helicobacter pylori
  ORGANISM  Helicobacter pylori
            Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
            Helicobacteraceae; Helicobacter.
REFERENCE   1  (bases 1 to 1119)
  AUTHORS   Wang,G., Rasko,D.A., Sherburne,R. and Taylor,D.E.
  TITLE     Molecular genetic basis for the variable expression of Lewis Y
            antigen in Helicobacter pylori: analysis of the alpha (1,2)
            fucosyltransferase gene
  JOURNAL   Mol. Microbiol. 31 (4), 1265-1274 (1999)
   PUBMED   10096092
REFERENCE   2  (bases 1 to 1119)
  AUTHORS   Wang,G. and Taylor,D.E.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JUL-1998) Medical Microbiology and Immunology,
            University of Alberta, 1-51 Medical Science Building, Edmonton, AB
            T6G 2H7, Canada
FEATURES             Location/Qualifiers
     source          1. .1119
                     /organism="Helicobacter pylori"
                     /mol_type="genomic DNA"
                     /strain="UA802"
                     /db_xref="taxon:210"
     gene            137. .1039
                     /gene="fucT2"
                     /allele="fucT2-type I"
     CDS             137. .1039
                     /gene="fucT2"
                     /allele="fucT2-type I"
                     /function="biosynthesis of Lewis Y antigen"
                     /note="a(1,2) fucosyltransferase"
                     /codon_start=1
                     /transl_table=11
                     /product="alpha-1,2-fucosyltransferase"
                     /protein_id="AAC99764.1"
                     /db_xref="GI:4093139"
                     /translation="MAFKVVQICGGLGNQMFQYAFAKSLQKHLNTPVLLDTTSFDWSN
                     RKMQLELFPIDLPYANAKEIAIAKMQHLPKLVRDALKYIGFDRVSQEIVFEYEPKLLK
                     PSRLTYFFGYFQDPRYFDAISSLIKQTFTLPPPPENNKNNNKKEEEYQRKLSLILAAK
                     NSVFVHIRRGDYVGIGCQLGIDYQKKALEYMAKRVPNMELFVFCEDLKFTQNLDLGYP
                     FTDMTTRDKEEEAYWDMLLMQSCKHGIIANSTYSWWAAYLMENPEKIIIGPKHWLFGH
                     ENILCKEWVKIESHFEVKSQKYNA"
ORIGIN

  Query Match           56.4%;  Score 22;  DB 1;  Length 1119;
  Best Local Similarity  100.0%;  Pred. No. 52;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         18 TTATAGCGAGCGTCTAATTCTA 39
              ||||||||||||||||||||||
Db       1098 TTATAGCGAGCGTCTAATTCTA 1119


RESULT 9
AR154400
LOCUS       AR154400               1119 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 1 from patent US 6238894.
ACCESSION   AR154400
VERSION     AR154400.1  GI:15122453
KEYWORDS    .
SOURCE      Unknown.
```

```
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1119)
  AUTHORS   Taylor,D., Wang,G. and Palcic,M.
  TITLE     .alpha.1,2 fucosyltransferase
  JOURNAL   Patent: US 6238894-A 1 29-MAY-2001;
FEATURES             Location/Qualifiers
     source          1. .1119
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             56.4%;  Score 22;  DB 6;  Length 1119;
  Best Local Similarity   100.0%;  Pred. No. 52;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         18 TTATAGCGAGCGTCTAATTCTA 39
              ||||||||||||||||||||||
Db       1098 TTATAGCGAGCGTCTAATTCTA 1119


RESULT 10
BD238089
LOCUS       BD238089               1119 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Helicobacter pylori-origin alpha-1,2-fucosyltransferase.
ACCESSION   BD238089
VERSION     BD238089.1  GI:33047859
KEYWORDS    JP 2002528122-A/1.
SOURCE      Helicobacter pylori
  ORGANISM  Helicobacter pylori
            Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
            Helicobacteraceae; Helicobacter.
REFERENCE   1  (bases 1 to 1119)
  AUTHORS   Taylor,D.E., Wang,G. and Palcic,M.
  TITLE     Helicobacter pylori-origin alpha-1,2-fucosyltransferase
  JOURNAL   Patent: JP 2002528122-A 1 03-SEP-2002;
            GOVERNORS OF THE UNIVERSITY OF ALBERTA
COMMENT     OS   Helicobacter pylori
            PN   JP 2002528122-A/1
            PD   03-SEP-2002
            PF   03-NOV-1999 JP 2000579755
            PR   04-NOV-1998 US    60/107268,02-NOV-1999 US    09/433598 PI
            DIANE E TAYLOR,GE WANG,MONICA PALCIC
            PC   C12N15/09,C07K16/40,C12N1/15,C12N1/19,C12N1/21,C12N5/10 PC
            ,C12N9/10,C12P19/18,
            PC   C12P21/08,C12Q1/48,C12Q1/68,G01N33/569//(C12N9/10,C12R1:01),
            PC   (C12N9/10,C12R1:19),(C12P19/18,C12R1:19),C12N15/00,C12N5/00 CC
              Helicobacter pylori-origin alpha-1,2-fucosyltransferase FH   Key
                        Location/Qualifiers
            FT   CDS             (137). .(1036).
FEATURES             Location/Qualifiers
     source          1. .1119
                     /organism="Helicobacter pylori"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:210"
ORIGIN

  Query Match             56.4%;  Score 22;  DB 6;  Length 1119;
  Best Local Similarity   100.0%;  Pred. No. 52;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         18 TTATAGCGAGCGTCTAATTCTA 39
              ||||||||||||||||||||||
Db       1098 TTATAGCGAGCGTCTAATTCTA 1119


RESULT 11
AR442784
LOCUS       AR442784               1119 bp    DNA     linear   PAT 20-FEB-2004
DEFINITION  Sequence 1 from patent US 6670160.
ACCESSION   AR442784
VERSION     AR442784.1  GI:42670204
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1119)
  AUTHORS   Taylor,D., Wang,G. and Palcic,M.
  TITLE     .alpha.1,2-fucosyltransferase
  JOURNAL   Patent: US 6670160-A 1 30-DEC-2003;
            Governors of the University of Alberta; Edmonton;
            CAX;
FEATURES             Location/Qualifiers
     source          1. .1119
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             56.4%;  Score 22;  DB 6;  Length 1119;
  Best Local Similarity   100.0%;  Pred. No. 52;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         18 TTATAGCGAGCGTCTAATTCTA 39
              ||||||||||||||||||||||
Db       1098 TTATAGCGAGCGTCTAATTCTA 1119


RESULT 12
AF093828
LOCUS       AF093828               1122 bp    DNA     linear   BCT 13-MAY-1999
DEFINITION  Helicobacter pylori strain UA1234 alpha-1,2-fucosyltransferase
            gene, complete cds.
ACCESSION   AF093828
VERSION     AF093828.1  GI:4808587
KEYWORDS    .
SOURCE      Helicobacter pylori
  ORGANISM  Helicobacter pylori
            Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
            Helicobacteraceae; Helicobacter.
REFERENCE   1  (bases 1 to 1122)
  AUTHORS   Wang,G., Rasko,D.A., Sherburne,R. and Taylor,D.E.
  TITLE     Molecular genetic basis for the variable expression of Lewis Y
            antigen in Helicobacter pylori: analysis of the alpha (1,2)
            fucosyltransferase gene
  JOURNAL   Mol. Microbiol. 31 (4), 1265-1274 (1999)
   PUBMED   10096092
REFERENCE   2  (bases 1 to 1122)
  AUTHORS   Wang,G. and Taylor,D.E.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-SEP-1998) Medical Microbiology and Immunology,
            University of Alberta, 1-51 Medical Science Building, Edmonton, AB
            T6G 2H7, Canada
FEATURES             Location/Qualifiers
     source          1. .1122
                     /organism="Helicobacter pylori"
                     /mol_type="genomic DNA"
                     /strain="UA1234"
                     /db_xref="taxon:210"
     CDS             142. .1041
                     /codon_start=1
                     /transl_table=11
                     /product="alpha-1,2-fucosyltransferase"
                     /protein_id="AAD29863.1"
                     /db_xref="GI:4808588"
                     /translation="MAFKVVQICGGLGNQMFQYAFAKSLQKHSNTPVLLDITSFDWSN
                     RKMQLELFPIDLPYASEKEIAIAKMQHLPKLVRNVLKCMGFDRVSQEIVFEYEPKLLK
                     TSRLTYFYGYFQDPRYFDAISPLIKQTFTLPPPPENGNNKKKEEEYHRKLALILAAKN
                     SVFVHIRRGDYVGIGCQLGIDYQKKALEYMAKRVPNMELFVFCEDLEFTQNLDLGYPF
                     MDMTTRDKEEEAYWDMLLMQSCKHGIIANSTYSWWAAYLINNPEKIIIGPKHWLFGHE
                     NILCKEWVKIESHFEVKSQKYNA"
ORIGIN

  Query Match             56.4%;  Score 22;  DB 1;  Length 1122;
  Best Local Similarity   100.0%;  Pred. No. 52;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy         18 TTATAGCGAGCGTCTAATTCTA 39
              ||||||||||||||||||||||
Db       1101 TTATAGCGAGCGTCTAATTCTA 1122


RESULT 13
AF093829
LOCUS       AF093829               1124 bp    DNA     linear   BCT 13-MAY-1999
DEFINITION  Helicobacter pylori strain UA1182 alpha-1,2-fucosyltransferase
            gene, alternative products, complete cds.
ACCESSION   AF093829
VERSION     AF093829.1  GI:4808589
KEYWORDS    .
SOURCE      Helicobacter pylori
  ORGANISM  Helicobacter pylori
            Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
            Helicobacteraceae; Helicobacter.
REFERENCE   1  (bases 1 to 1124)
  AUTHORS   Wang,G., Rasko,D.A., Sherburne,R. and Taylor,D.E.
  TITLE     Molecular genetic basis for the variable expression of Lewis Y
            antigen in Helicobacter pylori: analysis of the alpha (1,2)
            fucosyltransferase gene
  JOURNAL   Mol. Microbiol. 31 (4), 1265-1274 (1999)
   PUBMED   10096092
REFERENCE   2  (bases 1 to 1124)
  AUTHORS   Wang,G. and Taylor,D.E.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-SEP-1998) Medical Microbiology and Immunology,
            University of Alberta, 1-51 Medical Science Building, Edmonton, AB
            T6G 2H7, Canada
FEATURES             Location/Qualifiers
     source          1. .1124
                     /organism="Helicobacter pylori"
                     /mol_type="genomic DNA"
                     /strain="UA1182"
                     /db_xref="taxon:210"
     CDS             join(141. .573,573. .1042)
                     /note="ribosomal slippage"
                     /codon_start=1
                     /transl_table=11
                     /product="alpha-1,2-fucosyltransferase long form"
                     /protein_id="AAD29865.1"
                     /db_xref="GI:4808591"
                     /translation="MAFKVVRICGGLGNQMFQYAFAKSLQKHSNTPVLLDITSFDWSN
                     RKMQLELFPIDLPYASAKEIAIAKMQHLPKLVRDALKYIGFDRVSQEIVFEYEPKLLK
                     PSRLTYFYGYFQDPRYFDAISPLIKQTFTLPPPPKIIRIIIKKEEEYHRKLALILAAK
                     NSVFVHIRRGDYVGIGCQLGIDYQKKALEYMAKRVPNMELFVFCEDLTFTQNLDLGYP
                     FMDMTTRDKDEEAYWDMLLMQSCQHGIIANSTYSWWAAYLIENPEKIIIGPKHWLFGH
                     ENILCKEWVKIESHFEVKSQKYNA"
     CDS             141. .605
                     /note="nonfunctional"
                     /codon_start=1
                     /transl_table=11
                     /product="alpha-1,2-fucosyltransferase short form"
                     /protein_id="AAD29864.1"
                     /db_xref="GI:4808590"
                     /translation="MAFKVVRICGGLGNQMFQYAFAKSLQKHSNTPVLLDITSFDWSN
                     RKMQLELFPIDLPYASAKEIAIAKMQHLPKLVRDALKYIGFDRVSQEIVFEYEPKLLK
                     PSRLTYFYGYFQDPRYFDAISPLIKQTFTLPPPPKIIRIIIKKRKNTTANLL"
ORIGIN

  Query Match             56.4%;  Score 22;  DB 1;  Length 1124;
  Best Local Similarity   100.0%;  Pred. No. 52;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         18 TTATAGCGAGCGTCTAATTCTA 39
              ||||||||||||||||||||||
Db       1103 TTATAGCGAGCGTCTAATTCTA 1124


RESULT 14
AF093832
LOCUS       AF093832               1124 bp    DNA     linear   BCT 13-MAY-1999
DEFINITION  Helicobacter pylori strain UA1207 alpha-1,2-fucosyltransferase
            gene, complete cds.
ACCESSION   AF093832
VERSION     AF093832.1  GI:4808596
KEYWORDS    .
SOURCE      Helicobacter pylori
  ORGANISM  Helicobacter pylori
            Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
            Helicobacteraceae; Helicobacter.
REFERENCE   1  (bases 1 to 1124)
  AUTHORS   Wang,G., Rasko,D.A., Sherburne,R. and Taylor,D.E.
  TITLE     Molecular genetic basis for the variable expression of Lewis Y
            antigen in Helicobacter pylori: analysis of the alpha (1,2)
            fucosyltransferase gene
  JOURNAL   Mol. Microbiol. 31 (4), 1265-1274 (1999)
   PUBMED   10096092
REFERENCE   2  (bases 1 to 1124)
  AUTHORS   Wang,G. and Taylor,D.E.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-SEP-1998) Medical Microbiology and Immunology,
            University of Alberta, 1-51 Medical Science Building, Edmonton, AB
            T6G 2H7, Canada
FEATURES             Location/Qualifiers
     source          1. .1124
                     /organism="Helicobacter pylori"
                     /mol_type="genomic DNA"
                     /strain="UA1207"
                     /db_xref="taxon:210"
     CDS             143. .1042
                     /codon_start=1
                     /transl_table=11
                     /product="alpha-1,2-fucosyltransferase"
                     /protein_id="AAD29868.1"
                     /db_xref="GI:4808597"
                     /translation="MAFKVVQICGGLGNQMFQYAFAKSLQKHSNTPVLLDISSFDWSN
                     RKMQLELFPIDLPYASAKEIAIAKMQHLPKLVRDALKCMGFDRVSQEIVFEYEPKLLK
                     PSRLTYFYGYFQDPRYFDAISSLIKQTFTLPPPPENGNNKKKEEEYHRKLSLILAAQN
                     SVFVHIRRGDYVGIGCQLGIDYQKKALEYMAKRVPNMELFVFCEDLEFTQNLDLGYPF
                     MDMTTRNKEEEAYWDMLLMQSCKHGIIANSTYSWWAAYLINNPEKIIIGPKHWLFGHE
                     NILCKEWVKIESHFEVKSQKYNA"
ORIGIN

  Query Match             56.4%;  Score 22;  DB 1;  Length 1124;
  Best Local Similarity   100.0%;  Pred. No. 52;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         18 TTATAGCGAGCGTCTAATTCTA 39
              ||||||||||||||||||||||
Db       1103 TTATAGCGAGCGTCTAATTCTA 1124


RESULT 15
AF093833
LOCUS       AF093833               1124 bp    DNA     linear   BCT 13-MAY-1999
DEFINITION  Helicobacter pylori strain UA1210 alpha-1,2-fucosyltransferase
            gene, complete cds.
ACCESSION   AF093833
VERSION     AF093833.1  GI:4808598
KEYWORDS    .
SOURCE      Helicobacter pylori
  ORGANISM  Helicobacter pylori
            Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
            Helicobacteraceae; Helicobacter.
REFERENCE   1  (bases 1 to 1124)
  AUTHORS   Wang,G., Rasko,D.A., Sherburne,R. and Taylor,D.E.
  TITLE     Molecular genetic basis for the variable expression of Lewis Y
            antigen in Helicobacter pylori: analysis of the alpha (1,2)
            fucosyltransferase gene
  JOURNAL   Mol. Microbiol. 31 (4), 1265-1274 (1999)
   PUBMED   10096092
REFERENCE   2  (bases 1 to 1124)
  AUTHORS   Wang,G. and Taylor,D.E.
  TITLE     Direct Submission
```

```
  JOURNAL   Submitted (24-SEP-1998) Medical Microbiology and Immunology,
            University of Alberta, 1-51 Medical Science Building, Edmonton, AB
            T6G 2H7, Canada
FEATURES             Location/Qualifiers
     source          1. .1124
                     /organism="Helicobacter pylori"
                     /mol_type="genomic DNA"
                     /strain="UA1210"
                     /db_xref="taxon:210"
     CDS             143. .1045
                     /codon_start=1
                     /transl_table=11
                     /product="alpha-1,2-fucosyltransferase"
                     /protein_id="AAD29869.1"
                     /db_xref="GI:4808599"
                     /translation="MAFKVVQICGGLGNQMFQYAFAKSLQKHSNTPVLLDITSFDWSD
                     RKMQLELFPINLPYASAKEIAIAKMQHLPKLVRDALKCMGFDRVSQEIVFEYEPELLK
                     PSRLTYFYGYFQDPRYFDAISPLIKQTFTLPPPPENNKNNNKKEEEYHRKLSLILAAK
                     NSVFVHIRRGDYVGIGCQLGIDYQKKALEYMAKRVPNMELFVFCEDLEFTQNLDLGYP
                     FMDMTTRNKEEEAYWDMLLMQSCQHGIIANSTYSWWAAYLIENPEKIIIGPKHWLFGH
                     ENILCKEWVKIESHFEVKSQKYNA"
ORIGIN

  Query Match             56.4%;  Score 22;  DB 1;  Length 1124;
  Best Local Similarity   100.0%;  Pred. No. 52;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         18 TTATAGCGAGCGTCTAATTCTA 39
              ||||||||||||||||||||||
Db       1103 TTATAGCGAGCGTCTAATTCTA 1124


Search completed: January 30, 2006, 02:34:18
Job time : 898 secs
```

OM nucleic - nucleic search, using sw model

Run on: January 30, 2006, 01:24:06 ; Search time 1556 Seconds
(without alignments)
1172.684 Million cell updates/sec

Title: US-10-663-033-3-REVCOMP-4
Perfect score: 39
Sequence: 1 gaacactcacacgcgtcttatagcgagcgtctaattcta 39

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 0.1

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_htc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	23.8	61.0	1454	9	BZ554770	BZ554770 pacs1-60_
c	2	22.6	57.9	392	5	BY415162	BY415162 BY415162
	3	22.6	57.9	500	6	CD599349	CD599349 RK115A4H0
c	4	22	56.4	415	5	BY456864	BY456864 BY456864
c	5	22	56.4	599	2	BG606513	BG606513 WHE2957_B
	6	22	56.4	657	5	BQ805941	BQ805941 WHE3572_G
	7	22	56.4	664	1	AJ793428	AJ793428 AJ793428
c	8	22	56.4	680	5	BQ743392	BQ743392 WHE4103_D
c	9	22	56.4	700	6	CD866027	CD866027 AZO2.102G
	10	22	56.4	711	6	CD490533	CD490533 WHE2957_B
	11	22	56.4	719	1	AJ796733	AJ796733 AJ796733
c	12	22	56.4	728	2	BE972908	BE972908 601652552
	13	22	56.4	744	1	AJ797420	AJ797420 AJ797420
	14	22	56.4	762	1	AJ796856	AJ796856 AJ796856
c	15	22	56.4	836	8	CV761672	CV761672 FGAS05606
c	16	22	56.4	839	8	CV761673	CV761673 FGAS05606
	17	22	56.4	1007	11	CNS05ET8	AL334133 Tetraodon
	18	22	56.4	1116	7	CK207657	CK207657 FGAS01930
c	19	22	56.4	1124	8	DR736695	DR736695 FGAS08206
c	20	21.8	55.9	387	5	BY580486	BY580486 BY580486
c	21	21.8	55.9	399	5	BY506980	BY506980 BY506980
c	22	21.8	55.9	412	5	BY431624	BY431624 BY431624
c	23	21.8	55.9	449	8	M75881	M75881 CEL06A9 Chr
c	24	21.8	55.9	865	2	BF572390	BF572390 602076450
c	25	21.6	55.4	431	7	CV039590	CV039590 4136975 B
c	26	21.6	55.4	494	5	BU290064	BU290064 604162423
c	27	21.6	55.4	560	5	BU360261	BU360261 603476787
c	28	21.6	55.4	576	1	AJ735914	AJ735914 AJ735914
c	29	21.6	55.4	595	1	AJ450530	AJ450530 AJ450530
	30	21.6	55.4	623	1	AJ735924	AJ735924 AJ735924
c	31	21.6	55.4	630	9	BH764985	BH764985 BMBAC352H
c	32	21.6	55.4	647	3	BM427220	BM427220 pgf2n.pk0
c	33	21.6	55.4	685	7	CN233065	CN233065 WLB055F12
c	34	21.6	55.4	692	1	AJ452211	AJ452211 AJ452211
c	35	21.6	55.4	727	7	CK356301	CK356301 AGENCOURT
c	36	21.6	55.4	761	5	BU387220	BU387220 603861130
c	37	21.6	55.4	817	1	AJ450262	AJ450262 AJ450262
c	38	21.6	55.4	843	5	BU293792	BU293792 603577337
c	39	21.6	55.4	863	5	BU324884	BU324884 603489962
c	40	21.4	54.9	627	10	AG271419	AG271419 Cyanidios
	41	21.4	54.9	680	9	CC418927	CC418927 PUHJN76TB
c	42	21.4	54.9	680	9	CC418930	CC418930 PUHJN76TD
c	43	21.4	54.9	920	9	CC357981	CC357981 PUHRW41TB
	44	21.2	54.4	453	1	AA099175	AA099175 zk85e08.s
	45	21.2	54.4	741	10	BX157310	BX157310 Danio rer

ALIGNMENTS

RESULT 1
BZ554770
LOCUS BZ554770 1454 bp DNA linear GSS 17-DEC-2002
DEFINITION pacs1-60_4851.x1 pacs1-60 Pseudomonas aeruginosa genomic clone pacs1-60_4851, genomic survey sequence.
ACCESSION BZ554770
VERSION BZ554770.1 GI:27162705
KEYWORDS GSS.
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas.
REFERENCE 1 (bases 1 to 1454)
AUTHORS Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M., Burns,J.L., Kaul,R. and Olsen,M.V.
TITLE Whole-Genome-Sequence variation among multiple isolates of Psedomonas aeruginosa library
JOURNAL J. Bacteriol. (2002) In press
COMMENT Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.
FEATURES Location/Qualifiers
source 1. .1454
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="1-60"
/db_xref="taxon:287"
/clone="pacs1-60_4851"
/clone_lib="pacs1-60"
/note="clinical isolate 1-60 Whole genomic shotgun library."
ORIGIN

Query Match 61.0%; Score 23.8; DB 9; Length 1454;
Best Local Similarity 80.0%; Pred. No. 11;
Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
Qy          4 CACTCACACGCGTCTTATAGCGAGCGTCTAATTCT 38
              | ||||| ||||| |||| | |||||||| ||||
Db        906 CCCTCACTTGCGTCGTATATCTAGCGTCTACTTCT 940
```

RESULT 2
BY415162/c
LOCUS BY415162 392 bp mRNA linear EST 13-DEC-2002
DEFINITION BY415162 RIKEN full-length enriched, 17 days embryo kidney Mus musculus cDNA clone I920009E12 3', mRNA sequence.
ACCESSION BY415162
VERSION BY415162.1 GI:26680286
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 392)
AUTHORS Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S., Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I., Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A., Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C., Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H., Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V., Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A., Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T., Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S., Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A., Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A., Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R., Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T., Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G., Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S., Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M., Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K., Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M., Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y., Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I., Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P., Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M., Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K., Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S., Rogers,J., Birney,E. and Hayashizaki,Y.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
PUBMED 12466851
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hirozane,T., Imotani,K., Ishii,Y., Itoh,M., Kawai,J., Konno,H., Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R., Ohno,M., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y. Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.
FEATURES Location/Qualifiers
source 1. .392
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="I920009E12"
/tissue_type="kidney"
/dev_stage="17 days embryo"
/clone_lib="RIKEN full-length enriched, 17 days embryo kidney"
ORIGIN

Query Match 57.9%; Score 22.6; DB 5; Length 392;
Best Local Similarity 86.2%; Pred. No. 32;
Matches 25; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy          8 CACACGCGTCTTATAGCGAGCGTCTAATT 36
              |||||| | |||||||||||| || |||||
Db        162 CACACGGGACTTATAGCGAGAGTATAATT 134
```

RESULT 3
CD599349
LOCUS CD599349 500 bp mRNA linear EST 16-JUN-2003
DEFINITION RK115A4H02.T3 Zebrafish Kidney Marrow cDNA library Danio rerio cDNA clone RK115A4H02 5', mRNA sequence.
ACCESSION CD599349
VERSION CD599349.1 GI:31780225
KEYWORDS EST.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 500)
AUTHORS Song,H.D., Sun,X.J., Deng,M., Zhang,G.W., Zhou,Y., Wu,X.Y., Sheng,Y., Chen,Y., Ruan,Z., Jiang,C.L., Fan,H.Y., Zon,L.I., Kanki,J.P., Liu,T.X., Look,A.T. and Chen,Z.
TITLE Hematopoietic gene expression profile in zebrafish kidney marrow
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 101 (46), 16240-16245 (2004)
PUBMED 15520368
COMMENT Contact: Chen Z.
State Key Lab for Medical Genomics
Shanghai Institute of Hematology, Ruijin Hospital Affiliated to Shanghai Second Medical University
197 Rui Jin Road II, Shanghai 200025, P. R. China
Tel: 86-21-64740490
Fax: 86-21-64743206
Email: zchen@stn.sh.cn
Seq primer: T3.
FEATURES Location/Qualifiers
source 1. .500
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="RK115A4H02"
/dev_stage="mature"
/clone_lib="Zebrafish Kidney Marrow cDNA library"
/note="Organ: kidney; Vector: pBS-CMV; Site_1: XhoI; Site_2: EcoRI; Total RNA was extracted from the kidney tissues of mature zebrafish. The poly (A)+ RNA fraction was separated from total RNA by oligo (dT) cellulose chromatography. Library was initially constructed in the lambdaZAP Express vector (Stratagene) and in vivo excised into pBS-CMV vector."
ORIGIN

```
  Query Match             57.9%;  Score 22.6;  DB 6;  Length 500;
  Best Local Similarity   75.7%;  Pred. No. 33;
  Matches   28;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 GAACACTCACACGCGTCTTATAGCGAGCGTCTAATTC 37
              ||||||||  ||  |  ||  ||  |||||  |||||  |||
Db        165 GAACACTCCCAGGTTTCCTACAGCGAAGGTCTACTTC 201


RESULT 4
BY456864/c
LOCUS       BY456864                 415 bp    mRNA    linear   EST 13-DEC-2002
DEFINITION  BY456864 RIKEN full-length enriched, 17 days embryo whole body Mus
            musculus cDNA clone K630094C08 3', mRNA sequence.
ACCESSION   BY456864
VERSION     BY456864.1  GI:26755007
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 415)
  AUTHORS   Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
            Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I.,
            Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,
            Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,
            Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,
            Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V.,
            Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A.,
            Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,
            Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,
            Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,
            Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,
            Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,
            Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,
            Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G.,
            Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,
            Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,
            Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K.,
            Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,
            Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y.,
            Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,
            Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,
            Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,
            Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
            Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,
            Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,
            Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,
            Rogers,J., Birney,E. and Hayashizaki,Y.
  TITLE     Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
   PUBMED   12466851
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
            Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S.,
            Hirozane,T., Imotani,K., Ishii,Y., Itoh,M., Kawai,J., Konno,H.,
            Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R.,
            Ohno,M., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K.,
            Shiraki,T., Tagami,M., Waki,K., Watahiki,A., Muramatsu,M. and
            Hayashizaki,Y. Direct Submission
             Computational Analysis of Full-Length Mouse cDNAs Compared with
            Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. 10 (10), 1617-1630 (2000)
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res.
            10 (11), 1757-1771 (2000)
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
             cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
             Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
FEATURES              Location/Qualifiers
     source           1. .415
                      /organism="Mus musculus"
                      /mol_type="mRNA"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="K630094C08"
                      /tissue_type="whole body"
                      /dev_stage="17 days embryo"
                      /clone_lib="RIKEN full-length enriched, 17 days embryo
                      whole body"
ORIGIN

  Query Match             56.4%;  Score 22;  DB 5;  Length 415;
  Best Local Similarity   83.3%;  Pred. No. 61;
  Matches   25;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          8 CACACGCGTCTTATAGCGAGCGTCTAATTC 37
              ||||||  |  ||||||||||  ||  ||||||
Db        185 CACACGGGATTTATAGCGAGAGTATAATTC 156


RESULT 5
BG606513/c
LOCUS       BG606513                 599 bp    mRNA    linear   EST 17-APR-2001
DEFINITION  WHE2957_B01_D01ZS Wheat dormant embryo cDNA library Triticum
            aestivum cDNA clone WHE2957_B01_D01, mRNA sequence.
ACCESSION   BG606513
VERSION     BG606513.1  GI:13656484
KEYWORDS    EST.
SOURCE      Triticum aestivum (bread wheat)
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 599)
  AUTHORS   Anderson,O.D., Chao,S., Chin,A., Close,T.J., Doherty,L.,
            Fenton,R.D., Lazo,G.R., Rausch,C.J., Walker-Simmons,M.K. and
            Wilson,C.
  TITLE     The structure and function of the expressed portion of the wheat
            genomes - Dormant embryo cDNA library
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Olin Anderson
            US Department of Agriculture, Agriculture Research Service, Pacific
            West Area, Western Regional Research Center
            800 Buchanan Street, Albany, CA 94710, USA
            Tel: 5105595773
            Fax: 5105595818
            Email: oandersn@pw.usda.gov
            Sequence have been trimmed to remove vector sequence and low
            quality sequence with phred score less than 20
            Seq primer: Stratagene SK primer.
FEATURES              Location/Qualifiers
     source           1. .599
                      /organism="Triticum aestivum"
                      /mol_type="mRNA"
                      /cultivar="Brevor"
                      /db_xref="taxon:4565"
                      /clone="WHE2957_B01_D01"
```

/tissue_type="Seed embryo"
/dev_stage="Mature seed"
/lab_host="E. coli SOLR"
/clone_lib="Wheat dormant embryo cDNA library"
/note="Vector: Lambda Uni-ZAP XR, excised phagemid; Site_1: EcoRI; Site_2: XhoI; Plants were grown to seed maturity under conditions favoring seed dormancy (L. Dohery at K. Walker_Simmons lab, Washington State University, Pullman, WA). Embryos were cut from mature dormant seed (Doherty). Total RNA was prepared from these embryos, polyA was purified, a cDNA library was made, and the cDNA clones were in vivo excised to give pBluescript phagemids in the TJ Close lab at the University of California, Riverside (Chin, Fenton). Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."

ORIGIN

Query Match 56.4%; Score 22; DB 2; Length 599;
Best Local Similarity 73.7%; Pred. No. 63;
Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```
Qy          2 AACACTCACACGCGTCTTATAGCGAGCGTCTAATTCTA 39
              |||||||||  ||||||  ||  |  ||  |||  |  |||  |
Db        106 AACACTCACCTGCGTCTCATTGTGACCGTTTGCTTCAA 69
```

RESULT 6
BQ805941
LOCUS BQ805941 657 bp mRNA linear EST 31-JUL-2002
DEFINITION WHE3572_G04_M08ZS Wheat developing grains cDNA library Triticum aestivum cDNA clone WHE3572_G04_M08, mRNA sequence.
ACCESSION BQ805941
VERSION BQ805941.1 GI:22030150
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 657)
AUTHORS Altenbach,S., Anderson,O.D., Chao,S., Chin,A., Close,T.J., Cronin,K., Crossman,C., Fenton,R.D., Lazo,G.R., Pham,J., Rausch,C.J., Wilson,C. and Woo,J.
TITLE The structure and function of the expressed portion of the wheat genomes - Developing grains cDNA library
JOURNAL Unpublished (2002)
COMMENT Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: oandersn@pw.usda.gov
Sequences have been trimmed to remove vector sequence and low quality sequence with phred score less than 20
Seq primer: SK primer.
FEATURES Location/Qualifiers
source 1. .657
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Butte 86"
/db_xref="taxon:4565"
/clone="WHE3572_G04_M08"
/tissue_type="whole grains"
/dev_stage="3-44 days post anthesis seed"
/lab_host="E. coli SOLR"
/clone_lib="Wheat developing grains cDNA library"
/note="Vector: Lambda ZAP II, excised phagemid; Site_1: EcoRI; Plants were grown under six following different environmental regimes in greenhouse, Environment 1) 24oC/17oC day/night, well-watered, with post-anthesis fertilizer, Environment 2) 24oC/17oC day/night, well-watered, without post-anthesis fertilizer, Environment 3) 37oC/17oC day/night, well-watered, with post-anthesis fertilizer, Environment 4) 37oC/17oC day/night, well-watered, without post-anthesis fertilizer, Environment 5) 37oC/17oC day/night plus drought, with post-anthesis fertilizer, Environment 6) 37oC/17oC day/night plus drought, without post-anthesis fertilizer, developing wheat grains from the following were excised and frozen in liquid nitrogen, Environment 1 at 3, 5, 7, 8, 10, 12, 16, 20, 24, 28, 32, 36, 40, 44 DPA Environment 2 at 3, 5, 7, 8, 10, 12, 16, 20, 24, 28, 32, 36, 40, 44 DPA Environment 3 at 3, 5, 7, 8, 10, 12, 16, 20, 24, 28, 32, 34 DPA Environment 4 at 3, 5, 7, 8, 10, 12, 16, 20, 24, 28, 32, 34 DPA Environment 5 at 3, 5, 7, 8, 10, 12, 16, 20, 24, 28, 30 DPA Environment 6 at 3, 5, 7, 8, 10, 12, 16, 20, 24, 28, 30 DPA and total RNA was prepared by S. Altenbach and K. Cronin at USDA-ARS, Albany, CA. A cDNA library was made using poly (A) RNA, and the cDNA clones were in vivo excised to give pBluescript SK(-) phagemids in the TJ Close lab (Chin, Close, Fenton) at the University of California, Riverside. Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (others)."

ORIGIN

Query Match 56.4%; Score 22; DB 5; Length 657;
Best Local Similarity 73.7%; Pred. No. 64;
Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```
Qy          2 AACACTCACACGCGTCTTATAGCGAGCGTCTAATTCTA 39
              |||||||||  ||||||  ||  |  ||  |||  |  |||  |
Db        599 AACACTCACCTGCGTCTCATTGTGACCGTTTGCTTCAA 636
```

RESULT 7
AJ793428
LOCUS AJ793428 664 bp mRNA linear EST 08-DEC-2004
DEFINITION AJ793428 Antirrhinum majus whole plant Antirrhinum majus cDNA clone 018_3_02_f19, mRNA sequence.
ACCESSION AJ793428
VERSION AJ793428.1 GI:51108756
KEYWORDS EST.
SOURCE Antirrhinum majus (snapdragon)
ORGANISM Antirrhinum majus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; lamiids; Lamiales; Plantaginaceae; Antirrhineae; Antirrhinum.
REFERENCE 1 (bases 1 to 664)
AUTHORS Bey,M., Stueber,K., Fellenberg,K., Schwarz-Sommer,Z., Sommer,H., Saedler,H. and Zachgo,S.
TITLE Characterization of Antirrhinum Petal Development and Identification of Target Genes of the Class B MADS Box Gene DEFICIENS
JOURNAL Plant Cell 16 (12), 3197-3215 (2004)
PUBMED 15539471
COMMENT Contact: Schwarz-Sommer Z
Molekulare Pflanzengenetik
MPI fuer Zuechtungsforschung
Carl-von-Linne Weg 10, D-50829, Germany.
FEATURES Location/Qualifiers
source 1. .664
/organism="Antirrhinum majus"
/mol_type="mRNA"
/db_xref="taxon:4151"
/clone="018_3_02_f19"
/tissue_type="whole plant"
/clone_lib="Antirrhinum majus whole plant"

ORIGIN

Query Match 56.4%; Score 22; DB 1; Length 664;
Best Local Similarity 73.7%; Pred. No. 64;
Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```
Qy          1 GAACACTCACACGCGTCTTATAGCGAGCGTCTAATTCT 38
              || ||  |||||||||    || |||||||  ||||||
Db         54 GAGCAACCACACGCGTGCCGTATCGAGCGTGAAATTCT 91


RESULT 8
BQ743392/c
LOCUS       BQ743392                680 bp    mRNA    linear   EST 17-JUL-2002
DEFINITION  WHE4103_D12_H23ZS Wheat salt-stressed root cDNA library Triticum
            aestivum cDNA clone WHE4103_D12_H23, mRNA sequence.
ACCESSION   BQ743392
VERSION     BQ743392.1  GI:21890179
KEYWORDS    EST.
SOURCE      Triticum aestivum (bread wheat)
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 680)
  AUTHORS   Anderson,O.D., Akhunov,E., Chao,S., Crossman,C., Deal,K.,
            Dvorak,J., Lazo,G.R., Pham,J., Rausch,C.J., Wilson,C. and Woo,J.
  TITLE     The structure and function of the expressed portion of the wheat
            genomes - Salt-stressed root cDNA library
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Olin Anderson
            US Department of Agriculture, Agriculture Research Service, Pacific
            West Area, Western Regional Research Center
            800 Buchanan Street, Albany, CA 94710, USA
            Tel: 5105595773
            Fax: 5105595818
            Email: oandersn@pw.usda.gov
            Sequences have been trimmed to remove vector sequence and low
            quality sequence with phred score less than 20
            Seq primer: SK primer.
FEATURES             Location/Qualifiers
     source          1. .680
                     /organism="Triticum aestivum"
                     /mol_type="mRNA"
                     /cultivar="Chinese Spring"
                     /db_xref="taxon:4565"
                     /clone="WHE4103_D12_H23"
                     /tissue_type="Roots"
                     /dev_stage="Full tillering"
                     /lab_host="E. coli SOLR"
                     /clone_lib="Wheat salt-stressed root cDNA library"
                     /note="Vector: Lambda Uni-ZAP XR, excised phagemid
                     pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; Hydroponic
                     plants grown to full tillering stage were treated with 150
                     mM NaCl for either 12 hours or 7 days. Root tissues of the
                     plants subjected to both types of treatment were collected
                     separately at University of California, Davis (E. Akhunov
                     and K. Deal in J. Dvorak's Lab). Total RNA was prepared
                     separately from the two samples (12h and 7day treatments),
                     and equal amount of RNA was then pooled. PolyA RNA was
                     purified from the pooled RNA, a cDNA library was made, and
                     the cDNA clones were in vivo excised to give pBluescript
                     SK(-) phagemids in J. Dvorak's lab (E. Akhunov, J. Dvorak)
                     at the University of California, Davis. Colony plating,
                     plasmid DNA preparations and DNA sequencing were performed
                     in the OD Anderson lab (all other authors)."
ORIGIN

  Query Match             56.4%;  Score 22;  DB 5;  Length 680;
  Best Local Similarity   73.7%;  Pred. No. 64;
  Matches   28;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          2 AACACTCACACGCGTCTTATAGCGAGCGTCTAATTCTA 39
              |||||||||  |||||| || | || ||| |  ||| |
Db        172 AACACTCACCTGCGTCTCATTGTGACCGTTTGCTTCAA 135


RESULT 9
CD866027/c
LOCUS       CD866027                700 bp    mRNA    linear   EST 11-JUL-2003
DEFINITION  AZO2.102G05F001123 AZO2 Triticum aestivum cDNA clone AZO2102G05,
            mRNA sequence.
ACCESSION   CD866027
VERSION     CD866027.1  GI:32549843
KEYWORDS    EST.
SOURCE      Triticum aestivum (bread wheat)
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 700)
  AUTHORS   Genoplante.
  TITLE     Genoplante, a major partnership french program in plant genomics
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Genoplante
            Genoplante
            93, rue Henri Rochefort 91025 EVRY CEDEX France
            Tel: 33 1 69 47 54 00
            Fax: 33 1 69 47 54 10
            This sequence has been generated in the framework of the french
            plant genomics programme 'Genoplante' (http://www.genoplante.com
            and http://genoplante-info.infobiogen.fr).
FEATURES             Location/Qualifiers
     source          1. .700
                     /organism="Triticum aestivum"
                     /mol_type="mRNA"
                     /cultivar="recital"
                     /db_xref="taxon:4565"
                     /clone="AZO2102G05"
                     /tissue_type="root"
                     /clone_lib="AZO2"
ORIGIN

  Query Match             56.4%;  Score 22;  DB 6;  Length 700;
  Best Local Similarity   73.7%;  Pred. No. 64;
  Matches   28;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          2 AACACTCACACGCGTCTTATAGCGAGCGTCTAATTCTA 39
              |||||||||  |||||| || | || ||| |  ||| |
Db        163 AACACTCACCTGCGTCTCATTGTGACCGTTTGCTTCAA 126


RESULT 10
CD490533
LOCUS       CD490533                711 bp    mRNA    linear   EST 05-JUN-2003
DEFINITION  WHE2957_B01_D01ZT Brevor wheat dormant embryo cDNA library Triticum
            aestivum cDNA clone WHE2957_B01_D01, mRNA sequence.
ACCESSION   CD490533
VERSION     CD490533.1  GI:31417579
KEYWORDS    EST.
SOURCE      Triticum aestivum (bread wheat)
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 711)
  AUTHORS   Anderson,O.D., Chao,S., Chin,A., Close,T.J., Crossman,C.,
            Doherty,L., Fenton,R.D., Lazo,G.R., Rausch,C.J.,
            Walker-Simmons,M.K. and Wilson,C.
  TITLE     The structure and function of the expressed portion of the wheat
            genomes - Brevor dormant embryo cDNA library
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Olin Anderson
            US Department of Agriculture, Agriculture Research Service, Pacific
            West Area, Western Regional Research Center
            800 Buchanan Street, Albany, CA 94710, USA
            Tel: 5105595773
            Fax: 5105595818
            Email: oandersn@pw.usda.gov
            This EST was generated by sequencing from the 3' end of the clone.
            Sequences have been trimmed to remove vector sequence and low
```

```
          quality sequence with phred score less than 20.
          Seq primer: T7 primer.
FEATURES             Location/Qualifiers
     source          1. .711
                     /organism="Triticum aestivum"
                     /mol_type="mRNA"
                     /cultivar="Brevor"
                     /db_xref="taxon:4565"
                     /clone="WHE2957_B01_D01"
                     /tissue_type="Seed embryo"
                     /dev_stage="Mature seed"
                     /lab_host="E. coli SOLR"
                     /clone_lib="Brevor wheat dormant embryo cDNA library"
                     /note="Vector: Lambda Uni-ZAP XR, excised phagemid;
                     Site_1: EcoRI; Site_2: XhoI; Plants were grown to seed
                     maturity under conditions favoring seed dormancy (L.
                     Dohery at K. Walker_Simmons lab, Washington State
                     University, Pullman, WA). Embryos were cut from mature
                     dormant seed (Doherty). Total RNA was prepared from these
                     embryos, polyA was purified, a cDNA library was made, and
                     the cDNA clones were in vivo excised to give pBluescript
                     phagemids in the TJ Close lab at the University of
                     California, Riverside (Chin, Fenton). Plasmid DNA
                     preparations and DNA sequencing were performed in the OD
                     Anderson lab (all other authors)."
ORIGIN

  Query Match             56.4%;  Score 22;  DB 6;  Length 711;
  Best Local Similarity   73.7%;  Pred. No. 64;
  Matches   28;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          2 AACACTCACACGCGTCTTATAGCGAGCGTCTAATTCTA 39
              |||||||||  ||||||  || | ||  ||| |  ||| |
Db        608 AACACTCACCTGCGTCTCATTGTGACCGTTTGCTTCAA 645


RESULT 11
AJ796733
LOCUS       AJ796733                719 bp    mRNA    linear   EST 08-DEC-2004
DEFINITION  AJ796733 Antirrhinum majus whole plant Antirrhinum majus cDNA clone
            018_3_11_n14, mRNA sequence.
ACCESSION   AJ796733
VERSION     AJ796733.1  GI:51112061
KEYWORDS    EST.
SOURCE      Antirrhinum majus (snapdragon)
  ORGANISM  Antirrhinum majus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            asterids; lamiids; Lamiales; Plantaginaceae; Antirrhineae;
            Antirrhinum.
REFERENCE   1  (bases 1 to 719)
  AUTHORS   Bey,M., Stueber,K., Fellenberg,K., Schwarz-Sommer,Z., Sommer,H.,
            Saedler,H. and Zachgo,S.
  TITLE     Characterization of Antirrhinum Petal Development and
            Identification of Target Genes of the Class B MADS Box Gene
            DEFICIENS
  JOURNAL   Plant Cell 16 (12), 3197-3215 (2004)
   PUBMED   15539471
COMMENT     Contact: Schwarz-Sommer Z
            Molekulare Pflanzengenetik
            MPI fuer Zuechtungsforschung
            Carl-von-Linne Weg 10, D-50829, Germany.
FEATURES             Location/Qualifiers
     source          1. .719
                     /organism="Antirrhinum majus"
                     /mol_type="mRNA"
                     /db_xref="taxon:4151"
                     /clone="018_3_11_n14"
                     /tissue_type="whole plant"
                     /clone_lib="Antirrhinum majus whole plant"
ORIGIN

  Query Match             56.4%;  Score 22;  DB 1;  Length 719;
  Best Local Similarity   73.7%;  Pred. No. 64;
  Matches   28;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          1 GAACACTCACACGCGTCTTATAGCGAGCGTCTAATTCT 38
              || ||  |||||||||     || |||||||  ||||||
Db         61 GAGCAACCACACGCGTGCCGTATCGAGCGTGAAATTCT 98


RESULT 12
BE972908/c
LOCUS       BE972908                728 bp    mRNA    linear   EST 04-OCT-2000
DEFINITION  601652552R2 NIH_MGC_82 Homo sapiens cDNA clone IMAGE:3935676 3',
            mRNA sequence.
ACCESSION   BE972908
VERSION     BE972908.1  GI:10586244
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 728)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONETECH Laboratories, Inc.
             cDNA Library Preparation: CLONETECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM778  row: h  column: 13
            High quality sequence stop: 2.
FEATURES             Location/Qualifiers
     source          1. .728
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3935676"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_82"
                     /note="Organ: testis; Vector: pDNR-LIB (Clontech); Site_1:
                     SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc); 5' and
                     3' adaptors were used in cloning as follows: 5' adaptor
                     sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:
                     5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A,
                     C, or G and N = A, C,  G, or T).   Average insert size
                     1.35 kb (range 0.9-4.0 kb).  14/15 colonies contained
                     inserts by PCR. This library was enriched for full-length
                     clones and was constructed by Clontech  Laboratories (Palo
                     Alto, CA)."
ORIGIN

  Query Match             56.4%;  Score 22;  DB 2;  Length 728;
  Best Local Similarity   73.7%;  Pred. No. 64;
  Matches   28;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          1 GAACACTCACACGCGTCTTATAGCGAGCGTCTAATTCT 38
              |||||||||    || |||||  || ||| || |  ||| |
Db        357 GAACACTCATCTGCCTCTTAAAGAGAGAGTGTGATTAT 320


RESULT 13
AJ797420
LOCUS       AJ797420                744 bp    mRNA    linear   EST 08-DEC-2004
DEFINITION  AJ797420 Antirrhinum majus whole plant Antirrhinum majus cDNA clone
            018_4_01_m23, mRNA sequence.
ACCESSION   AJ797420
VERSION     AJ797420.1  GI:51112748
KEYWORDS    EST.
```

```
SOURCE      Antirrhinum majus (snapdragon)
  ORGANISM  Antirrhinum majus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            asterids; lamiids; Lamiales; Plantaginaceae; Antirrhineae;
            Antirrhinum.
REFERENCE   1  (bases 1 to 744)
  AUTHORS   Bey,M., Stueber,K., Fellenberg,K., Schwarz-Sommer,Z., Sommer,H.,
            Saedler,H. and Zachgo,S.
  TITLE     Characterization of Antirrhinum Petal Development and
            Identification of Target Genes of the Class B MADS Box Gene
            DEFICIENS
  JOURNAL   Plant Cell 16 (12), 3197-3215 (2004)
   PUBMED   15539471
COMMENT     Contact: Schwarz-Sommer Z
            Molekulare Pflanzengenetik
            MPI fuer Zuechtungsforschung
            Carl-von-Linne Weg 10, D-50829, Germany.
FEATURES             Location/Qualifiers
     source          1. .744
                     /organism="Antirrhinum majus"
                     /mol_type="mRNA"
                     /db_xref="taxon:4151"
                     /clone="018_4_01_m23"
                     /tissue_type="whole plant"
                     /clone_lib="Antirrhinum majus whole plant"
ORIGIN

  Query Match             56.4%;  Score 22;  DB 1;  Length 744;
  Best Local Similarity   73.7%;  Pred. No. 65;
  Matches   28;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          1 GAACACTCACACGCGTCTTATAGCGAGCGTCTAATTCT 38
              || ||  |||||||||     || |||||||  ||||||
Db         62 GAGCAACCACACGCGTGCCGTATCGAGCGTGAAATTCT 99


RESULT 14
AJ796856
LOCUS       AJ796856                762 bp    mRNA    linear   EST 08-DEC-2004
DEFINITION  AJ796856 Antirrhinum majus whole plant Antirrhinum majus cDNA clone
            018_3_12_c22, mRNA sequence.
ACCESSION   AJ796856
VERSION     AJ796856.1  GI:51112184
KEYWORDS    EST.
SOURCE      Antirrhinum majus (snapdragon)
  ORGANISM  Antirrhinum majus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            asterids; lamiids; Lamiales; Plantaginaceae; Antirrhineae;
            Antirrhinum.
REFERENCE   1  (bases 1 to 762)
  AUTHORS   Bey,M., Stueber,K., Fellenberg,K., Schwarz-Sommer,Z., Sommer,H.,
            Saedler,H. and Zachgo,S.
  TITLE     Characterization of Antirrhinum Petal Development and
            Identification of Target Genes of the Class B MADS Box Gene
            DEFICIENS
  JOURNAL   Plant Cell 16 (12), 3197-3215 (2004)
   PUBMED   15539471
COMMENT     Contact: Schwarz-Sommer Z
            Molekulare Pflanzengenetik
            MPI fuer Zuechtungsforschung
            Carl-von-Linne Weg 10, D-50829, Germany.
FEATURES             Location/Qualifiers
     source          1. .762
                     /organism="Antirrhinum majus"
                     /mol_type="mRNA"
                     /db_xref="taxon:4151"
                     /clone="018_3_12_c22"
                     /tissue_type="whole plant"
                     /clone_lib="Antirrhinum majus whole plant"
ORIGIN

  Query Match             56.4%;  Score 22;  DB 1;  Length 762;
  Best Local Similarity   73.7%;  Pred. No. 65;
  Matches   28;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          1 GAACACTCACACGCGTCTTATAGCGAGCGTCTAATTCT 38
              || ||  |||||||||     || |||||||  ||||||
Db         82 GAGCAACCACACGCGTGCCGTATCGAGCGTGAAATTCT 119


RESULT 15
CV761672/c
LOCUS       CV761672                836 bp    mRNA    linear   EST 09-NOV-2004
DEFINITION  FGAS056060 Triticum aestivum FGAS: Library 2 Gate 3 Triticum
            aestivum cDNA, mRNA sequence.
ACCESSION   CV761672
VERSION     CV761672.1  GI:55602063
KEYWORDS    EST.
SOURCE      Triticum aestivum (bread wheat)
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 836)
  AUTHORS   Allard,F., Crosby,W.L., Danyluk,J., Eudes,F., Frick,M., Gaudet,D.,
            Genswein,B., Graf,R., Gulick,P., Hrycan,L.D., Laroche,A.,
            Links,M.G., McCarthy,B.L., Monroy,A., Muzak,I., Nilson,D.,
            Penniket,C., Roach,J.L. and Sarhan,F.
  TITLE     Functional Genomics of Abiotic Stress In Wheat and Canola Crops
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Wm L Crosby
            Bioinformatics
            University of Saskatchewan, Department of Computer Science
            1C101 Engineering Building, 57 Campus Drive, Saskatoon,
            Saskatchewan, S7N 5A9, Canada
            Tel: 306 966 1769
            Fax: 306 966 2033
            Email: fgas_ests@cs.usask.ca
            This sequence is the direct result of the Base calling software
            Phred (default parameters). It is the raw base calls. To aid in the
            identification of the high quality insert the software Lucy
            (default parameters) has been run on this sequence. Lucy identified
            the region [95,721].
            Plate: WEF027  row: M  column: 22.
FEATURES             Location/Qualifiers
     source          1. .836
                     /organism="Triticum aestivum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4565"
                     /clone_lib="Triticum aestivum FGAS: Library 2 Gate 3"
                     /note="Organ: Crown and leaf; Vector: pCMV.SPORT6; Aerial
                     parts (crown and leaf) of wheat cultivar Norstar from
                     control and long exposure times to low temperature. 4 mRNA
                     populations were combined before constructing the library;
                     7 days non-acclimated plants and 1, 23, and 53 days
                     cold-acclimated at 4C. Non-acclimated and cold-acclimated
                     plants were grown in vermiculite This is the only library
                     that was done according to the Invitrogen manual, and
                     therefore, a percentage of clones will not have the 3
                     prime end because of NotI digestion within the cDNA."
ORIGIN

  Query Match             56.4%;  Score 22;  DB 8;  Length 836;
  Best Local Similarity   73.7%;  Pred. No. 65;
  Matches   28;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          2 AACACTCACACGCGTCTTATAGCGAGCGTCTAATTCTA 39
              |||||||||  ||||||  || | ||  ||| |  ||| |
Db        277 AACACTCACCTGCGTCTCATTGTGACCGTTTGCTTCAA 240


Search completed: January 30, 2006, 03:26:16
Job time : 1559 secs
```

OM nucleic - nucleic search, using sw model

Run on: January 30, 2006, 01:49:51 ; Search time 73 Seconds
(without alignments)
949.656 Million cell updates/sec

Title: US-10-663-033-3-REVCOMP-4
Perfect score: 39
Sequence: 1 gaacactcacacgcgtcttatagcgagcgtctaattcta 39

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 0.1

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/1_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP_COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE_COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	22	56.4	1119	3	US-09-433-598-1	Sequence 1, Appli
	2	22	56.4	1119	3	US-09-848-838A-1	Sequence 1, Appli
c	3	20	51.3	20	3	US-09-433-598-4	Sequence 4, Appli
c	4	20	51.3	20	3	US-09-848-838A-4	Sequence 4, Appli
c	5	19.2	49.2	2664	3	US-09-489-039A-4547	Sequence 4547, Ap
	6	19.2	49.2	36899	3	US-09-949-016-13566	Sequence 13566, A
	7	19	48.7	19	3	US-09-433-598-3	Sequence 3, Appli
	8	19	48.7	19	3	US-09-848-838A-3	Sequence 3, Appli
	9	19	48.7	2670	3	US-09-489-039A-70	Sequence 70, Appl
c	10	19	48.7	124884	3	US-09-661-596A-76	Sequence 76, Appl
c	11	19	48.7	124884	3	US-09-913-514-1	Sequence 1, Appli
c	12	19	48.7	124884	3	US-10-288-823-76	Sequence 76, Appl
c	13	19	48.7	125157	3	US-09-913-514-2	Sequence 2, Appli
	14	18.8	48.2	537	3	US-09-710-279-1175	Sequence 1175, Ap
	15	18.8	48.2	563	3	US-09-495-050A-210	Sequence 210, App
c	16	18.8	48.2	601	3	US-09-949-002-3268	Sequence 3268, Ap
c	17	18.8	48.2	601	3	US-09-949-002-3269	Sequence 3269, Ap
c	18	18.8	48.2	601	3	US-09-949-002-10786	Sequence 10786, A
c	19	18.8	48.2	601	3	US-09-949-002-10787	Sequence 10787, A
	20	18.8	48.2	730	3	US-09-533-559-868	Sequence 868, App
c	21	18.8	48.2	2226	3	US-09-710-279-1177	Sequence 1177, Ap
c	22	18.8	48.2	2301	3	US-09-134-001C-852	Sequence 852, App
	23	18.8	48.2	3511	3	US-09-710-279-3461	Sequence 3461, Ap
	24	18.8	48.2	3705	3	US-09-710-279-3592	Sequence 3592, Ap
	25	18.8	48.2	34725	3	US-09-949-002-665	Sequence 665, App
	26	18.8	48.2	34726	3	US-09-949-002-857	Sequence 857, App
c	27	18.8	48.2	35688	3	US-09-949-016-16873	Sequence 16873, A
	28	18.8	48.2	49795	3	US-09-453-702B-60	Sequence 60, Appl
	29	18.8	48.2	49795	3	US-10-114-170-60	Sequence 60, Appl
	30	18.8	48.2	236341	3	US-09-949-016-13978	Sequence 13978, A
	31	18.8	48.2	786431	3	US-09-751-389-3	Sequence 3, Appli
	32	18.6	47.7	601	3	US-09-949-016-135264	Sequence 135264,
	33	18.6	47.7	107421	3	US-09-949-016-15532	Sequence 15532, A
c	34	18.4	47.2	601	3	US-09-949-016-168281	Sequence 168281,
	35	18.4	47.2	1397	3	US-09-270-767-5652	Sequence 5652, Ap
	36	18.4	47.2	1397	3	US-09-270-767-20934	Sequence 20934, A
c	37	18.4	47.2	2394	3	US-09-861-451A-11	Sequence 11, Appl
	38	18.4	47.2	2661	3	US-10-104-047-38	Sequence 38, Appl
c	39	18.4	47.2	60589	3	US-09-949-016-17070	Sequence 17070, A
c	40	18.4	47.2	116652	3	US-09-949-016-13413	Sequence 13413, A
c	41	18.4	47.2	225127	3	US-09-949-016-16480	Sequence 16480, A
	42	18.2	46.7	601	3	US-09-949-016-24751	Sequence 24751, A
	43	18.2	46.7	601	3	US-09-949-016-35121	Sequence 35121, A
c	44	18.2	46.7	1188	3	US-09-090-044B-1	Sequence 1, Appli
c	45	18.2	46.7	1323	3	US-09-533-559-5815	Sequence 5815, Ap

ALIGNMENTS

RESULT 1
US-09-433-598-1
; Sequence 1, Application US/09433598
; Patent No. 6238894
; GENERAL INFORMATION:
; APPLICANT: Taylor Dr., Diane
; APPLICANT: Wang, Ge
; APPLICANT: Palcic, Monica
; TITLE OF INVENTION: Alpha1, 2 Fucosyltransferase
; FILE REFERENCE: 07254/061001
; CURRENT APPLICATION NUMBER: US/09/433,598
; CURRENT FILING DATE: 1998-11-02
; EARLIER APPLICATION NUMBER: 60/107268
; EARLIER FILING DATE: 1998-11-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1119
; TYPE: DNA
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (137)..(1036)
US-09-433-598-1

Query Match 56.4%; Score 22; DB 3; Length 1119;
Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          18 TTATAGCGAGCGTCTAATTCTA 39
               ||||||||||||||||||||||
Db        1098 TTATAGCGAGCGTCTAATTCTA 1119
```

RESULT 2
US-09-848-838A-1
; Sequence 1, Application US/09848838A
; Patent No. 6670160
; GENERAL INFORMATION:
; APPLICANT: Taylor, Diane
; APPLICANT: Wang, Ge
; APPLICANT: Palcic, Monica
; TITLE OF INVENTION: ALPHA 1,2-FUCOSYLTRANSFERASE
; FILE REFERENCE: 07254-061002
; CURRENT APPLICATION NUMBER: US/09/848,838A
; CURRENT FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: US 09/433,598

```
;  PRIOR FILING DATE: 1999-11-02
;  PRIOR APPLICATION NUMBER: US 60/107,268
;  PRIOR FILING DATE: 1998-11-04
;  NUMBER OF SEQ ID NOS: 23
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 1119
;   TYPE: DNA
;   ORGANISM: Helicobacter pylori
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (137)...(1036)
US-09-848-838A-1

  Query Match             56.4%;  Score 22;  DB 3;  Length 1119;
  Best Local Similarity   100.0%;  Pred. No. 2.4;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         18 TTATAGCGAGCGTCTAATTCTA 39
              ||||||||||||||||||||||
Db       1098 TTATAGCGAGCGTCTAATTCTA 1119


RESULT 3
US-09-433-598-4/c
; Sequence 4, Application US/09433598
; Patent No. 6238894
; GENERAL INFORMATION:
;  APPLICANT: Taylor Dr., Diane
;  APPLICANT: Wang, Ge
;  APPLICANT: Palcic, Monica
;  TITLE OF INVENTION: Alpha1, 2 Fucosyltransferase
;  FILE REFERENCE: 07254/061001
;  CURRENT APPLICATION NUMBER: US/09/433,598
;  CURRENT FILING DATE: 1998-11-02
;  EARLIER APPLICATION NUMBER: 60/107268
;  EARLIER FILING DATE: 1998-11-04
;  NUMBER OF SEQ ID NOS: 23
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
;   LENGTH: 20
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:
;   OTHER INFORMATION: oligonucleotide
US-09-433-598-4

  Query Match             51.3%;  Score 20;  DB 3;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 9.4;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         20 ATAGCGAGCGTCTAATTCTA 39
              ||||||||||||||||||||
Db         20 ATAGCGAGCGTCTAATTCTA 1


RESULT 4
US-09-848-838A-4/c
; Sequence 4, Application US/09848838A
; Patent No. 6670160
; GENERAL INFORMATION:
;  APPLICANT: Taylor, Diane
;  APPLICANT:  Wang, Ge
;  APPLICANT:  Palcic, Monica
;  TITLE OF INVENTION: ALPHA 1,2-FUCOSYLTRANSFERASE
;  FILE REFERENCE: 07254-061002
;  CURRENT APPLICATION NUMBER: US/09/848,838A
;  CURRENT FILING DATE:  2002-05-21
;  PRIOR APPLICATION NUMBER: US 09/433,598
;  PRIOR FILING DATE: 1999-11-02
;  PRIOR APPLICATION NUMBER: US 60/107,268
;  PRIOR FILING DATE: 1998-11-04
;  NUMBER OF SEQ ID NOS: 23
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;   LENGTH: 20
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: synthetically generated oligonucleotide
US-09-848-838A-4

  Query Match             51.3%;  Score 20;  DB 3;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 9.4;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         20 ATAGCGAGCGTCTAATTCTA 39
              ||||||||||||||||||||
Db         20 ATAGCGAGCGTCTAATTCTA 1


RESULT 5
US-09-489-039A-4547/c
; Sequence 4547, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
;  APPLICANT:  Gary Breton et. al
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
;  TITLE OF INVENTION:  PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  2709.2004001
;  CURRENT APPLICATION NUMBER: US/09/489,039A
;  CURRENT FILING DATE:  2000-01-27
;  PRIOR APPLICATION NUMBER:  US 60/117,747
;  PRIOR FILING DATE:  1999-01-29
;  NUMBER OF SEQ ID NOS:  14342
; SEQ ID NO 4547
;   LENGTH: 2664
;   TYPE: DNA
;   ORGANISM: Klebsiella pneumoniae
US-09-489-039A-4547

  Query Match             49.2%;  Score 19.2;  DB 3;  Length 2664;
  Best Local Similarity   75.0%;  Pred. No. 55;
  Matches   24;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          8 CACACGCGTCTTATAGCGAGCGTCTAATTCTA 39
              || |||| ||| |||||| ||| |  |||| |
Db        375 CAGACGCATCTGATAGCGCGCGGCGCATTCGA 344


RESULT 6
US-09-949-016-13566
; Sequence 13566, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13566
;   LENGTH: 36899
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
```

```
;   LOCATION: (1)...(36899)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13566

  Query Match            49.2%;  Score 19.2;  DB 3;  Length 36899;
  Best Local Similarity  75.0%;  Pred. No. 90;
  Matches   24;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          3 ACACTCACACGCGTCTTATAGCGAGCGTCTAA 34
              ||||||||||    |||||||  ||| | |||
Db      11789 ACACTCACACATAACTTATAGGTAGCATTTAA 11820


RESULT 7
US-09-433-598-3
; Sequence 3, Application US/09433598
; Patent No. 6238894
; GENERAL INFORMATION:
;  APPLICANT: Taylor Dr., Diane
;  APPLICANT: Wang, Ge
;  APPLICANT: Palcic, Monica
;  TITLE OF INVENTION: Alpha1, 2 Fucosyltransferase
;  FILE REFERENCE: 07254/061001
;  CURRENT APPLICATION NUMBER: US/09/433,598
;  CURRENT FILING DATE: 1998-11-02
;  EARLIER APPLICATION NUMBER: 60/107268
;  EARLIER FILING DATE: 1998-11-04
;  NUMBER OF SEQ ID NOS: 23
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 19
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:
;   OTHER INFORMATION: oligonucleotide
US-09-433-598-3

  Query Match            48.7%;  Score 19;  DB 3;  Length 19;
  Best Local Similarity  100.0%;  Pred. No. 27;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAACACTCACACGCGTCTT 19
              |||||||||||||||||||
Db          1 GAACACTCACACGCGTCTT 19


RESULT 8
US-09-848-838A-3
; Sequence 3, Application US/09848838A
; Patent No. 6670160
; GENERAL INFORMATION:
;  APPLICANT: Taylor, Diane
;  APPLICANT:  Wang, Ge
;  APPLICANT:  Palcic, Monica
;  TITLE OF INVENTION: ALPHA 1,2-FUCOSYLTRANSFERASE
;  FILE REFERENCE: 07254-061002
;  CURRENT APPLICATION NUMBER: US/09/848,838A
;  CURRENT FILING DATE:  2002-05-21
;  PRIOR APPLICATION NUMBER: US 09/433,598
;  PRIOR FILING DATE: 1999-11-02
;  PRIOR APPLICATION NUMBER: US 60/107,268
;  PRIOR FILING DATE: 1998-11-04
;  NUMBER OF SEQ ID NOS: 23
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 19
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: synthetically generated oligonucleotide
US-09-848-838A-3

  Query Match            48.7%;  Score 19;  DB 3;  Length 19;
  Best Local Similarity  100.0%;  Pred. No. 27;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAACACTCACACGCGTCTT 19
              |||||||||||||||||||
Db          1 GAACACTCACACGCGTCTT 19


RESULT 9
US-09-489-039A-70
; Sequence 70, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
;  APPLICANT:  Gary Breton et. al
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
;  TITLE OF INVENTION:  PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  2709.2004001
;  CURRENT APPLICATION NUMBER: US/09/489,039A
;  CURRENT FILING DATE:  2000-01-27
;  PRIOR APPLICATION NUMBER:  US 60/117,747
;  PRIOR FILING DATE:  1999-01-29
;  NUMBER OF SEQ ID NOS:  14342
; SEQ ID NO 70
;   LENGTH: 2670
;   TYPE: DNA
;   ORGANISM: Klebsiella pneumoniae
US-09-489-039A-70

  Query Match            48.7%;  Score 19;  DB 3;  Length 2670;
  Best Local Similarity  71.4%;  Pred. No. 68;
  Matches   25;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          3 ACACTCACACGCGTCTTATAGCGAGCGTCTAATTC 37
              |||||  ||||  |||||  |||   ||| |||||
Db       1295 ACACTGGCACGATTCTTAATGCGTTAGTCGAATTC 1329


RESULT 10
US-09-661-596A-76/c
; Sequence 76, Application US/09661596A
; Patent No. 6528066
; GENERAL INFORMATION:
;  APPLICANT: Grose, Charles
;  APPLICANT:  Santos, Richard
;  TITLE OF INVENTION: VARIANT VARICELLA-ZOSTER VIRUSES AND METHODS OF USE
;  FILE REFERENCE: 140.0011 0101
;  CURRENT APPLICATION NUMBER: US/09/661,596A
;  CURRENT FILING DATE:  2000-09-14
;  PRIOR APPLICATION NUMBER: US 60/153,779
;  PRIOR FILING DATE: 1999-09-14
;  NUMBER OF SEQ ID NOS: 80
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 76
;   LENGTH: 124884
;   TYPE: DNA
;   ORGANISM: Varicella zoster
US-09-661-596A-76

  Query Match            48.7%;  Score 19;  DB 3;  Length 124884;
  Best Local Similarity  71.4%;  Pred. No. 1.4e+02;
  Matches   25;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          2 AACACTCACACGCGTCTTATAGCGAGCGTCTAATT 36
              || || |||||  | | | || ||| ||| |||||
Db       2483 AAAACACACACCAGACGTGTACCGAACGTTTAATT 2449


RESULT 11
US-09-913-514-1/c
; Sequence 1, Application US/09913514
; Patent No. 6653069
; GENERAL INFORMATION:
```

```
;  APPLICANT: GOMI, Yasuyuki
;  APPLICANT:  SUNAMACHI, Hiroki
;  APPLICANT:  TAKAHASHI, Michiaki
;  APPLICANT:  YAMANISHI, Koichi
;  TITLE OF INVENTION: Method for Quality Control of an Attenuated Varicella Live Vaccin
;  FILE REFERENCE: 0216-0454P
;  CURRENT APPLICATION NUMBER: US/09/913,514
;  CURRENT FILING DATE:  2001-12-07
;  PRIOR APPLICATION NUMBER: PCT/JP01/00678
;  PRIOR FILING DATE: 2001-01-31
;  PRIOR APPLICATION NUMBER: JP 2000-62734
;  PRIOR FILING DATE: 2000-01-31
;  NUMBER OF SEQ ID NOS: 42
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 124884
;   TYPE: DNA
;   ORGANISM: Varicella virus
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(124884)
;   OTHER INFORMATION: Dumas Strain
US-09-913-514-1

  Query Match            48.7%;  Score 19;  DB 3;  Length 124884;
  Best Local Similarity  71.4%;  Pred. No. 1.4e+02;
  Matches   25;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          2 AACACTCACACGCGTCTTATAGCGAGCGTCTAATT 36
              || || |||||  | | | || ||| ||| |||||
Db       2483 AAAACACACACCAGACGTGTACCGAACGTTTAATT 2449


RESULT 12
US-10-288-823-76/c
; Sequence 76, Application US/10288823
; Patent No. 6843997
; GENERAL INFORMATION:
;  APPLICANT: Grose, Charles
;  APPLICANT:  Santos, Richard
;  TITLE OF INVENTION: VARIANT VARICELLA-ZOSTER VIRUSES AND METHODS OF USE
;  FILE REFERENCE: 140.00110102
;  CURRENT APPLICATION NUMBER: US/10/288,823
;  CURRENT FILING DATE:  2002-11-06
;  PRIOR APPLICATION NUMBER: US 09/661,596
;  PRIOR FILING DATE: 2000-09-14
;  PRIOR APPLICATION NUMBER: US 60/153,779
;  PRIOR FILING DATE: 1999-09-14
;  NUMBER OF SEQ ID NOS: 80
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 76
;   LENGTH: 124884
;   TYPE: DNA
;   ORGANISM: Varicella zoster
US-10-288-823-76

  Query Match            48.7%;  Score 19;  DB 3;  Length 124884;
  Best Local Similarity  71.4%;  Pred. No. 1.4e+02;
  Matches   25;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          2 AACACTCACACGCGTCTTATAGCGAGCGTCTAATT 36
              || || |||||  | | | || ||| ||| |||||
Db       2483 AAAACACACACCAGACGTGTACCGAACGTTTAATT 2449


RESULT 13
US-09-913-514-2/c
; Sequence 2, Application US/09913514
; Patent No. 6653069
; GENERAL INFORMATION:
;  APPLICANT: GOMI, Yasuyuki
;  APPLICANT:  SUNAMACHI, Hiroki
;  APPLICANT:  TAKAHASHI, Michiaki
;  APPLICANT:  YAMANISHI, Koichi
;  TITLE OF INVENTION: Method for Quality Control of an Attenuated Varicella Live Vaccin
;  FILE REFERENCE: 0216-0454P
;  CURRENT APPLICATION NUMBER: US/09/913,514
;  CURRENT FILING DATE:  2001-12-07
;  PRIOR APPLICATION NUMBER: PCT/JP01/00678
;  PRIOR FILING DATE: 2001-01-31
;  PRIOR APPLICATION NUMBER: JP 2000-62734
;  PRIOR FILING DATE: 2000-01-31
;  NUMBER OF SEQ ID NOS: 42
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 125157
;   TYPE: DNA
;   ORGANISM: Varicella virus
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(125157)
;   OTHER INFORMATION: Attenuated Oka strain
US-09-913-514-2

  Query Match            48.7%;  Score 19;  DB 3;  Length 125157;
  Best Local Similarity  71.4%;  Pred. No. 1.4e+02;
  Matches   25;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          2 AACACTCACACGCGTCTTATAGCGAGCGTCTAATT 36
              || || |||||  | | | || ||| ||| |||||
Db       2482 AAAACACACACCAGACGTGTACCGAACGTTTAATT 2448


RESULT 14
US-09-710-279-1175
; Sequence 1175, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
;  APPLICANT: KIMMERLY, WILLIAM JOHN
;  TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
;  FILE REFERENCE: PU3480US
;  CURRENT APPLICATION NUMBER: US/09/710,279
;  CURRENT FILING DATE:  2000-11-09
;  PRIOR APPLICATION NUMBER: 60/164,258
;  PRIOR FILING DATE: 1999-11-09
;  NUMBER OF SEQ ID NOS: 4472
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1175
;   LENGTH: 537
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: synthetic
;   OTHER INFORMATION: nucleic acid sequence
US-09-710-279-1175

  Query Match            48.2%;  Score 18.8;  DB 3;  Length 537;
  Best Local Similarity  76.7%;  Pred. No. 62;
  Matches   23;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          9 ACACGCGTCTTATAGCGAGCGTCTAATTCT 38
              | |||| | ||  |||| ||||| ||||||
Db        375 ATACGCTTGTTTAAGCGTGCGTCCAATTCT 404


RESULT 15
US-09-495-050A-210
; Sequence 210, Application US/09495050A
; Patent No. 6492505
; GENERAL INFORMATION:
;  APPLICANT: Roopa, Reddy
;  APPLICANT:  Guegler, Karl, J.
;  APPLICANT:  Au-Young, Janice
;  TITLE OF INVENTION: COMPOSITION FOR DETECTION OF GENES ENCODING MEMBRANE-ASSOCIATED P
;  FILE REFERENCE: PA-0013 US
;  CURRENT APPLICATION NUMBER: US/09/495,050A
```

```
;  CURRENT FILING DATE:  2000-01-31
;  PRIOR APPLICATION NUMBER: 60/118,318
;  PRIOR FILING DATE: February 1, 1999
;  NUMBER OF SEQ ID NOS: 305
;  SOFTWARE: PERL Program
; SEQ ID NO 210
;   LENGTH: 563
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. 6492505 2378406CT1
US-09-495-050A-210

  Query Match             48.2%;  Score 18.8;  DB 3;  Length 563;
  Best Local Similarity   68.4%;  Pred. No. 62;
  Matches   26;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          2 AACACTCACACGCGTCTTATAGCGAGCGTCTAATTCTA 39
              || | | |||   ||||| |||| |  | ||||| |||
Db        313 AAAAGTAACATTAGTCTTCTAGCTAATGACTAATGCTA 350


Search completed: January 30, 2006, 03:28:46
Job time : 74 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: January 30, 2006, 02:34:27 ; Search time 349 Seconds
(without alignments)
924.085 Million cell updates/sec

Title: US-10-663-033-3-REVCOMP-4
Perfect score: 39
Sequence: 1 gaacactcacacgcgtcttatagcgagcgtctaattcta 39

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 0.1

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA_Main:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	22	56.4	934	3	US-09-895-913A-253	Sequence 253, App
	2	22	56.4	1119	3	US-09-848-838-1	Sequence 1, Appli
	3	22	56.4	1119	7	US-10-663-033-1	Sequence 1, Appli
	4	21.8	55.9	2295	5	US-10-128-714-2164	Sequence 2164, Ap
	5	21.8	55.9	2450	5	US-10-128-714-1164	Sequence 1164, Ap
	6	21.8	55.9	2505	5	US-10-128-714-6164	Sequence 6164, Ap
	7	21.8	55.9	2505	5	US-10-128-714-7164	Sequence 7164, Ap
	8	21.8	55.9	4450	5	US-10-128-714-164	Sequence 164, App
	9	21.8	55.9	4505	5	US-10-128-714-5164	Sequence 5164, Ap
c	10	20.6	52.8	852	5	US-10-027-632-120434	Sequence 120434,
c	11	20.6	52.8	852	5	US-10-027-632-120435	Sequence 120435,
c	12	20.6	52.8	852	6	US-10-027-632-120434	Sequence 120434,
c	13	20.6	52.8	852	6	US-10-027-632-120435	Sequence 120435,
	14	20.2	51.8	262	8	US-10-674-124A-12533	Sequence 12533, A
	15	20.2	51.8	331	8	US-10-674-124A-12532	Sequence 12532, A
c	16	20.2	51.8	2454	7	US-10-282-122A-36364	Sequence 36364, A
c	17	20.2	51.8	29973	3	US-09-984-429-262	Sequence 262, App
	18	20.2	51.8	29973	6	US-10-158-034-99	Sequence 99, Appl
c	19	20	51.3	20	3	US-09-848-838-4	Sequence 4, Appli
c	20	20	51.3	20	7	US-10-663-033-4	Sequence 4, Appli
c	21	20	51.3	117754	5	US-10-087-192-28	Sequence 28, Appl
	22	19.8	50.8	519	4	US-09-925-065A-415522	Sequence 415522,
c	23	19.8	50.8	2349	9	US-10-795-159-694	Sequence 694, App
	24	19.8	50.8	3469	9	US-10-795-159-449	Sequence 449, App
	25	19.8	50.8	191996	9	US-10-795-159-683	Sequence 683, App
c	26	19.8	50.8	263852	8	US-10-812-232-6	Sequence 6, Appli
c	27	19.6	50.3	600	9	US-10-972-079-66530	Sequence 66530, A
c	28	19.6	50.3	600	9	US-10-972-079-66531	Sequence 66531, A
c	29	19.6	50.3	189158	5	US-10-087-192-415	Sequence 415, App
c	30	19.6	50.3	192673	8	US-10-331-053-1	Sequence 1, Appli
c	31	19.4	49.7	546	5	US-10-027-632-215033	Sequence 215033,
c	32	19.4	49.7	546	6	US-10-027-632-215033	Sequence 215033,
	33	19.4	49.7	560	4	US-09-925-065A-438373	Sequence 438373,
	34	19.4	49.7	560	4	US-09-925-065A-438374	Sequence 438374,
c	35	19.4	49.7	576	7	US-10-332-859-290	Sequence 290, App
	36	19.4	49.7	599	4	US-09-925-065A-584840	Sequence 584840,
	37	19.4	49.7	618	5	US-10-027-632-179694	Sequence 179694,
	38	19.4	49.7	618	6	US-10-027-632-179694	Sequence 179694,
	39	19.4	49.7	627	5	US-10-027-632-291443	Sequence 291443,
	40	19.4	49.7	627	6	US-10-027-632-291443	Sequence 291443,
	41	19.4	49.7	653	3	US-09-988-067B-69	Sequence 69, Appl
	42	19.4	49.7	685	3	US-09-930-213-103	Sequence 103, App
	43	19.4	49.7	1695	4	US-09-925-065A-719835	Sequence 719835,
c	44	19.4	49.7	2450	6	US-10-388-934-599	Sequence 599, App
c	45	19.4	49.7	5758	3	US-09-866-050A-605	Sequence 605, App

ALIGNMENTS

RESULT 1
US-09-895-913A-253/c
; Sequence 253, Application US/09895913A
; Patent No. US20020160456A1
; GENERAL INFORMATION:
; APPLICANT: Kleanthous, Harold
; APPLICANT: Al-Garawi, Amal
; APPLICANT: Miller, Charles
; APPLICANT: Tomb, Jean Francois
; APPLICANT: Oomen, Raymond P.
; TITLE OF INVENTION: Identification of Polynucleotides
; TITLE OF INVENTION: Encoding No. US20020160456A1el Helicobacter Polypeptides in the
; TITLE OF INVENTION: Genome
; FILE REFERENCE: 06132/043002
; CURRENT APPLICATION NUMBER: US/09/895,913A
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 08/881,227
; PRIOR FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 253
; LENGTH: 934
; TYPE: DNA
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (51)...(881)
US-09-895-913A-253

Query Match 56.4%; Score 22; DB 3; Length 934;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         18 TTATAGCGAGCGTCTAATTCTA 39
              ||||||||||||||||||||||
Db        844 TTATAGCGAGCGTCTAATTCTA 823
```

RESULT 2
US-09-848-838-1
; Sequence 1, Application US/09848838
; Patent No. US20020037570A1
; GENERAL INFORMATION:
; APPLICANT: Taylor, Diane
; APPLICANT: Wang, Ge
; APPLICANT: Palcic, Monica

```
;  TITLE OF INVENTION: ALPHA 1,2-FUCOSYLTRANSFERASE
;  FILE REFERENCE: 07254-061002
;  CURRENT APPLICATION NUMBER: US/09/848,838
;  CURRENT FILING DATE:  2001-05-03
;  PRIOR APPLICATION NUMBER: US 09/433,598
;  PRIOR FILING DATE: 1999-11-02
;  PRIOR APPLICATION NUMBER: US 60/107,268
;  PRIOR FILING DATE: 1998-11-04
;  NUMBER OF SEQ ID NOS: 23
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 1119
;   TYPE: DNA
;   ORGANISM: Helicobacter pylori
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (137)...(1036)
US-09-848-838-1

  Query Match            56.4%;  Score 22;  DB 3;  Length 1119;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         18 TTATAGCGAGCGTCTAATTCTA 39
              ||||||||||||||||||||||
Db       1098 TTATAGCGAGCGTCTAATTCTA 1119


RESULT 3
US-10-663-033-1
; Sequence 1, Application US/10663033
; Publication No. US20040048331A1
; GENERAL INFORMATION:
;  APPLICANT: Taylor, Diane
;  APPLICANT:  Wang, Ge
;  APPLICANT:  Palcic, Monica
;  TITLE OF INVENTION: ALPHA 1,2-FUCOSYLTRANSFERASE
;  FILE REFERENCE: 07254-061002
;  CURRENT APPLICATION NUMBER: US/10/663,033
;  CURRENT FILING DATE:  2003-09-15
;  PRIOR APPLICATION NUMBER: US/09/848,838A
;  PRIOR FILING DATE: 2002-05-21
;  PRIOR APPLICATION NUMBER: US 09/433,598
;  PRIOR FILING DATE: 1999-11-02
;  PRIOR APPLICATION NUMBER: US 60/107,268
;  PRIOR FILING DATE: 1998-11-04
;  NUMBER OF SEQ ID NOS: 23
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 1119
;   TYPE: DNA
;   ORGANISM: Helicobacter pylori
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (137)...(1036)
US-10-663-033-1

  Query Match            56.4%;  Score 22;  DB 7;  Length 1119;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         18 TTATAGCGAGCGTCTAATTCTA 39
              ||||||||||||||||||||||
Db       1098 TTATAGCGAGCGTCTAATTCTA 1119


RESULT 4
US-10-128-714-2164
; Sequence 2164, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Bo
;  APPLICANT:  Hu, Wenqi
;  APPLICANT:  Tishkoff, Daniel
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Eroshkin, Alexey M
;  APPLICANT:  Lemieux, Sebastien M
;  TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
;  TITLE OF INVENTION:  Methods of Use
;  FILE REFERENCE: 10182-018-999
;  CURRENT APPLICATION NUMBER: US/10/128,714
;  CURRENT FILING DATE:  2002-04-23
;  PRIOR APPLICATION NUMBER: US 60/285,697
;  PRIOR FILING DATE: 2001-04-23
;  PRIOR APPLICATION NUMBER: US 60/287,066
;  PRIOR FILING DATE: 2001-04-27
;  PRIOR APPLICATION NUMBER: US 60/295,890
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: US 60/303,899
;  PRIOR FILING DATE: 2001-07-09
;  PRIOR APPLICATION NUMBER: US 60/316,362
;  PRIOR FILING DATE: 2001-08-31
;  NUMBER OF SEQ ID NOS: 8603
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2164
;   LENGTH: 2295
;   TYPE: DNA
;   ORGANISM: Aspergillus fumigatus
US-10-128-714-2164

  Query Match            55.9%;  Score 21.8;  DB 5;  Length 2295;
  Best Local Similarity   78.8%;  Pred. No. 19;
  Matches   26;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          3 ACACTCACACGCGTCTTATAGCGAGCGTCTAAT 35
              |||| || |||||||||| ||||||    ||||
Db       1546 ACACGCAAACGCGTCTTAGAGCGAGGAGATAAT 1578


RESULT 5
US-10-128-714-1164
; Sequence 1164, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Bo
;  APPLICANT:  Hu, Wenqi
;  APPLICANT:  Tishkoff, Daniel
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Eroshkin, Alexey M
;  APPLICANT:  Lemieux, Sebastien M
;  TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
;  TITLE OF INVENTION:  Methods of Use
;  FILE REFERENCE: 10182-018-999
;  CURRENT APPLICATION NUMBER: US/10/128,714
;  CURRENT FILING DATE:  2002-04-23
;  PRIOR APPLICATION NUMBER: US 60/285,697
;  PRIOR FILING DATE: 2001-04-23
;  PRIOR APPLICATION NUMBER: US 60/287,066
;  PRIOR FILING DATE: 2001-04-27
;  PRIOR APPLICATION NUMBER: US 60/295,890
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: US 60/303,899
;  PRIOR FILING DATE: 2001-07-09
;  PRIOR APPLICATION NUMBER: US 60/316,362
;  PRIOR FILING DATE: 2001-08-31
;  NUMBER OF SEQ ID NOS: 8603
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1164
;   LENGTH: 2450
;   TYPE: DNA
;   ORGANISM: Aspergillus fumigatus
US-10-128-714-1164

  Query Match            55.9%;  Score 21.8;  DB 5;  Length 2450;
  Best Local Similarity   78.8%;  Pred. No. 19;
  Matches   26;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;
```

```
Qy          3 ACACTCACACGCGTCTTATAGCGAGCGTCTAAT 35
              |||| || |||||||||| ||||||    ||||
Db       1546 ACACGCAAACGCGTCTTAGAGCGAGGAGATAAT 1578
```

RESULT 6
US-10-128-714-6164
; Sequence 6164, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wenqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroshkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; TITLE OF INVENTION: Methods of Use
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6164
; LENGTH: 2505
; TYPE: DNA
; ORGANISM: Aspergillus fumigatus
US-10-128-714-6164

Query Match 55.9%; Score 21.8; DB 5; Length 2505;
Best Local Similarity 78.8%; Pred. No. 19;
Matches 26; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
Qy          3 ACACTCACACGCGTCTTATAGCGAGCGTCTAAT 35
              |||| || |||||||||| ||||||    ||||
Db       1471 ACACGCAAACGCGTCTTAGAGCGAGGAGATAAT 1503
```

RESULT 7
US-10-128-714-7164
; Sequence 7164, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wenqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroshkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; TITLE OF INVENTION: Methods of Use
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7164
; LENGTH: 2505
; TYPE: DNA
; ORGANISM: Aspergillus fumigatus
US-10-128-714-7164

Query Match 55.9%; Score 21.8; DB 5; Length 2505;
Best Local Similarity 78.8%; Pred. No. 19;
Matches 26; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
Qy          3 ACACTCACACGCGTCTTATAGCGAGCGTCTAAT 35
              |||| || |||||||||| ||||||    ||||
Db       1471 ACACGCAAACGCGTCTTAGAGCGAGGAGATAAT 1503
```

RESULT 8
US-10-128-714-164
; Sequence 164, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wenqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroshkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; TITLE OF INVENTION: Methods of Use
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 164
; LENGTH: 4450
; TYPE: DNA
; ORGANISM: Aspergillus fumigatus
US-10-128-714-164

Query Match 55.9%; Score 21.8; DB 5; Length 4450;
Best Local Similarity 78.8%; Pred. No. 21;
Matches 26; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
Qy          3 ACACTCACACGCGTCTTATAGCGAGCGTCTAAT 35
              |||| || |||||||||| ||||||    ||||
Db       2546 ACACGCAAACGCGTCTTAGAGCGAGGAGATAAT 2578
```

RESULT 9
US-10-128-714-5164
; Sequence 5164, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wenqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroshkin, Alexey M
; APPLICANT: Lemieux, Sebastien M

```
;  TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
;  TITLE OF INVENTION:  Methods of Use
;  FILE REFERENCE: 10182-018-999
;  CURRENT APPLICATION NUMBER: US/10/128,714
;  CURRENT FILING DATE:  2002-04-23
;  PRIOR APPLICATION NUMBER: US 60/285,697
;  PRIOR FILING DATE: 2001-04-23
;  PRIOR APPLICATION NUMBER: US 60/287,066
;  PRIOR FILING DATE: 2001-04-27
;  PRIOR APPLICATION NUMBER: US 60/295,890
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: US 60/303,899
;  PRIOR FILING DATE: 2001-07-09
;  PRIOR APPLICATION NUMBER: US 60/316,362
;  PRIOR FILING DATE: 2001-08-31
;  NUMBER OF SEQ ID NOS: 8603
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5164
;   LENGTH: 4505
;   TYPE: DNA
;   ORGANISM: Aspergillus fumigatus
US-10-128-714-5164

  Query Match             55.9%;  Score 21.8;  DB 5;  Length 4505;
  Best Local Similarity   78.8%;  Pred. No. 21;
  Matches   26;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          3 ACACTCACACGCGTCTTATAGCGAGCGTCTAAT 35
              |||| || |||||||||| ||||||    ||||
Db       2471 ACACGCAAACGCGTCTTAGAGCGAGGAGATAAT 2503


RESULT 10
US-10-027-632-120434/c
; Sequence 120434, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 120434
;   LENGTH: 852
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-120434

  Query Match             52.8%;  Score 20.6;  DB 5;  Length 852;
  Best Local Similarity   74.3%;  Pred. No. 55;
  Matches   26;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          5 ACTCACACGCGTCTTATAGCGAGCGTCTAATTCTA 39
              | ||| || ||  ||||| |||  ||||||||||
Db         50 AATCATACTCGATTTATATGGAGACTCTAATTCTA 16


RESULT 11
US-10-027-632-120435/c
; Sequence 120435, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 120435
;   LENGTH: 852
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-120435

  Query Match             52.8%;  Score 20.6;  DB 5;  Length 852;
  Best Local Similarity   74.3%;  Pred. No. 55;
  Matches   26;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          5 ACTCACACGCGTCTTATAGCGAGCGTCTAATTCTA 39
              | ||| || ||  ||||| |||  ||||||||||
Db         50 AATCATACTCGATTTATATGGAGACTCTAATTCTA 16


RESULT 12
US-10-027-632-120434/c
; Sequence 120434, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 120434
;   LENGTH: 852
;   TYPE: DNA
```

```
;   ORGANISM: Human
US-10-027-632-120434

  Query Match             52.8%;  Score 20.6;  DB 6;  Length 852;
  Best Local Similarity   74.3%;  Pred. No. 55;
  Matches   26;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          5 ACTCACACGCGTCTTATAGCGAGCGTCTAATTCTA 39
              | ||| || ||  |||||  |||  ||||||||||
Db         50 AATCATACTCGATTTATATGGAGACTCTAATTCTA 16


RESULT 13
US-10-027-632-120435/c
; Sequence 120435, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 120435
;   LENGTH: 852
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-120435

  Query Match             52.8%;  Score 20.6;  DB 6;  Length 852;
  Best Local Similarity   74.3%;  Pred. No. 55;
  Matches   26;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          5 ACTCACACGCGTCTTATAGCGAGCGTCTAATTCTA 39
              | ||| || ||  |||||  |||  ||||||||||
Db         50 AATCATACTCGATTTATATGGAGACTCTAATTCTA 16


RESULT 14
US-10-674-124A-12533
; Sequence 12533, Application US/10674124A
; Publication No. US20040197797A1
; GENERAL INFORMATION:
;  APPLICANT: INOKO, Hidetoshi
;  APPLICANT:  TAMIYA, Gen
;  TITLE OF INVENTION: GENE MAPPING METHOD USING MICROSATELLITE
;  TITLE OF INVENTION:  GENETIC POLYMORPHISM MARKERS
;  FILE REFERENCE: ORIN-003CIP
;  CURRENT APPLICATION NUMBER: US/10/674,124A
;  CURRENT FILING DATE:  2003-09-26
;  PRIOR APPLICATION NUMBER: 10/257,511
;  PRIOR FILING DATE: 2003-03-07
;  PRIOR APPLICATION NUMBER: PCT/JP00/07621
;  PRIOR FILING DATE: 2000-10-30
;  PRIOR APPLICATION NUMBER: JP2000-112699
;  PRIOR FILING DATE: 2000-04-13
;  PRIOR APPLICATION NUMBER: JP2002-327516
;  PRIOR FILING DATE: 2002-09-28
;  PRIOR APPLICATION NUMBER: JP2002-383869
;  PRIOR FILING DATE: 2002-12-09
;  NUMBER OF SEQ ID NOS: 27110
; SEQ ID NO 12533
;   LENGTH: 262
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: AC000119.1_83235
;   FEATURE:
;   OTHER INFORMATION: Located on chromosome 7
;   FEATURE:
;   OTHER INFORMATION: Distance between a terminus base of telomere on
;   OTHER INFORMATION: chromosomal short arm and 5'-terminus of this base
;   OTHER INFORMATION: sequence : 94422145
;   FEATURE:
;   OTHER INFORMATION: Distance between 3'-terminus of neighbour sequence of
;   OTHER INFORMATION: sequence listing upward to telomere on chrosomal short arm and
;   OTHER INFORMATION: 5'-terminus of this base sequence : -331
US-10-674-124A-12533

  Query Match             51.8%;  Score 20.2;  DB 8;  Length 262;
  Best Local Similarity   75.8%;  Pred. No. 67;
  Matches   25;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          3 ACACTCACACGCGTCTTATAGCGAGCGTCTAAT 35
              |||| ||||| |||||||||  ||| | |  ||
Db         52 ACACACACACACGTCTTATATAGAGTGACATAT 84


RESULT 15
US-10-674-124A-12532
; Sequence 12532, Application US/10674124A
; Publication No. US20040197797A1
; GENERAL INFORMATION:
;  APPLICANT: INOKO, Hidetoshi
;  APPLICANT:  TAMIYA, Gen
;  TITLE OF INVENTION: GENE MAPPING METHOD USING MICROSATELLITE
;  TITLE OF INVENTION:  GENETIC POLYMORPHISM MARKERS
;  FILE REFERENCE: ORIN-003CIP
;  CURRENT APPLICATION NUMBER: US/10/674,124A
;  CURRENT FILING DATE:  2003-09-26
;  PRIOR APPLICATION NUMBER: 10/257,511
;  PRIOR FILING DATE: 2003-03-07
;  PRIOR APPLICATION NUMBER: PCT/JP00/07621
;  PRIOR FILING DATE: 2000-10-30
;  PRIOR APPLICATION NUMBER: JP2000-112699
;  PRIOR FILING DATE: 2000-04-13
;  PRIOR APPLICATION NUMBER: JP2002-327516
;  PRIOR FILING DATE: 2002-09-28
;  PRIOR APPLICATION NUMBER: JP2002-383869
;  PRIOR FILING DATE: 2002-12-09
;  NUMBER OF SEQ ID NOS: 27110
; SEQ ID NO 12532
;   LENGTH: 331
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: AC000119.1_83407
;   FEATURE:
;   OTHER INFORMATION: Located on chromosome 7
;   FEATURE:
;   OTHER INFORMATION: Distance between a terminus base of telomere on
;   OTHER INFORMATION: chromosomal short arm and 5'-terminus of this base
;   OTHER INFORMATION: sequence : 94422145
;   FEATURE:
;   OTHER INFORMATION: Distance between 3'-terminus of neighbour sequence of
;   OTHER INFORMATION: sequence listing upward to telomere on chrosomal short arm and
;   OTHER INFORMATION: 5'-terminus of this base sequence : 572
US-10-674-124A-12532

  Query Match             51.8%;  Score 20.2;  DB 8;  Length 331;
```

Best Local Similarity 75.8%; Pred. No. 70;
Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```
Qy          3 ACACTCACACGCGTCTTATAGCGAGCGTCTAAT 35
              |||| ||||| |||||||||  ||| | |  ||
Db         52 ACACACACACACGTCTTATATAGAGTGACATAT 84
```

Search completed: January 30, 2006, 05:23:21
Job time : 349 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: January 30, 2006, 03:26:26 ; Search time 392.5 Seconds
(without alignments)
82.538 Million cell updates/sec

Title: US-10-663-033-3-REVCOMP-4
Perfect score: 39
Sequence: 1 gaacactcacacgcgtcttatagcgagcgtctaattcta 39

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 0.1

Searched: 6059551 seqs, 415333918 residues

Total number of hits satisfying chosen parameters: 12119102

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA_New:*
1: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq1:*
8: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq2:*
10: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq3:*
11: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	20	51.3	953	7	US-10-750-185-47026	Sequence 47026, A
	2	20	51.3	953	7	US-10-750-623-47026	Sequence 47026, A
c	3	19.6	50.3	1036	7	US-10-750-185-53947	Sequence 53947, A
c	4	19.6	50.3	1036	7	US-10-750-623-53947	Sequence 53947, A
c	5	19.4	49.7	2424	7	US-10-750-185-26886	Sequence 26886, A
c	6	19.4	49.7	2424	7	US-10-750-623-26886	Sequence 26886, A
c	7	19.4	49.7	3518	8	US-11-136-527-485	Sequence 485, App
	8	19	48.7	732	7	US-10-750-185-53265	Sequence 53265, A
	9	19	48.7	732	7	US-10-750-623-53265	Sequence 53265, A
	10	18.8	48.2	201	7	US-10-995-561-27542	Sequence 27542, A
	11	18.8	48.2	201	7	US-10-995-561-27569	Sequence 27569, A
	12	18.8	48.2	537	7	US-10-793-626-1175	Sequence 1175, Ap
	13	18.8	48.2	648	7	US-10-750-185-42748	Sequence 42748, A
	14	18.8	48.2	648	7	US-10-750-623-42748	Sequence 42748, A
c	15	18.8	48.2	2226	7	US-10-793-626-1177	Sequence 1177, Ap
	16	18.8	48.2	3511	7	US-10-793-626-3461	Sequence 3461, Ap
	17	18.8	48.2	3705	7	US-10-793-626-3592	Sequence 3592, Ap
	18	18.8	48.2	43436	7	US-10-995-561-13240	Sequence 13240, A
c	19	18.6	47.7	1738	7	US-10-750-185-49884	Sequence 49884, A
c	20	18.6	47.7	1738	7	US-10-750-623-49884	Sequence 49884, A
c	21	18.6	47.7	1741	7	US-10-750-185-34115	Sequence 34115, A
c	22	18.6	47.7	1741	7	US-10-750-623-34115	Sequence 34115, A
	23	18.4	47.2	600	8	US-11-136-527-6622	Sequence 6622, Ap
	24	18.4	47.2	2802	8	US-11-136-527-2526	Sequence 2526, Ap
c	25	18.2	46.7	1853	7	US-10-750-185-51151	Sequence 51151, A
c	26	18.2	46.7	1853	7	US-10-750-623-51151	Sequence 51151, A
	27	18.2	46.7	168656	8	US-11-112-908-59	Sequence 59, Appl
	28	18.2	46.7	170285	8	US-11-112-908-58	Sequence 58, Appl
c	29	17.8	45.6	600	7	US-10-972-233-5	Sequence 5, Appli
c	30	17.8	45.6	1332	7	US-10-750-185-47157	Sequence 47157, A
c	31	17.8	45.6	1332	7	US-10-750-623-47157	Sequence 47157, A
c	32	17.8	45.6	2367	7	US-10-750-185-58552	Sequence 58552, A
c	33	17.8	45.6	2367	7	US-10-750-623-58552	Sequence 58552, A
c	34	17.6	45.1	1545	7	US-10-750-185-43979	Sequence 43979, A
c	35	17.6	45.1	1545	7	US-10-750-623-43979	Sequence 43979, A
	36	17.6	45.1	46271	7	US-10-995-561-13347	Sequence 13347, A
c	37	17.6	45.1	403278	7	US-10-995-561-13421	Sequence 13421, A
	38	17.4	44.6	600	8	US-11-136-527-7552	Sequence 7552, Ap
	39	17.4	44.6	1240	7	US-10-750-185-26047	Sequence 26047, A
	40	17.4	44.6	1240	7	US-10-750-623-26047	Sequence 26047, A
	41	17.4	44.6	1613	8	US-11-024-959-71	Sequence 71, Appl
c	42	17.4	44.6	2143	7	US-10-750-185-60653	Sequence 60653, A
c	43	17.4	44.6	2143	7	US-10-750-623-60653	Sequence 60653, A
	44	17.4	44.6	2507	8	US-11-136-527-3456	Sequence 3456, Ap
c	45	17.4	44.6	187786	7	US-10-995-561-13474	Sequence 13474, A

ALIGNMENTS

RESULT 1
US-10-750-185-47026
; Sequence 47026, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DeNISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MMI1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 47026
; LENGTH: 953
; TYPE: DNA
; ORGANISM: Bovine 19866880944977
US-10-750-185-47026

Query Match 51.3%; Score 20; DB 7; Length 953;
Best Local Similarity 72.2%; Pred. No. 7.2;
Matches 26; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```
Qy          3 ACACTCACACGCGTCTTATAGCGAGCGTCTAATTCT 38
              | ||||| ||  |||||||||   |  ||||||| ||
Db        133 ATACTCATACATGTCTTATACTCAATGTCTAATACT 168
```

RESULT 2
US-10-750-623-47026
; Sequence 47026, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DeNISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom

```
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-1
;  CURRENT APPLICATION NUMBER: US/10/750,623
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 47026
;   LENGTH: 953
;   TYPE: DNA
;   ORGANISM: Bovine  19866880944977
US-10-750-623-47026

  Query Match             51.3%;  Score 20;  DB 7;  Length 953;
  Best Local Similarity   72.2%;  Pred. No. 7.2;
  Matches   26;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          3 ACACTCACACGCGTCTTATAGCGAGCGTCTAATTCT 38
              | |||||  ||  ||||||||   |  ||||||| ||
Db        133 ATACTCATACATGTCTTATACTCAATGTCTAATACT 168


RESULT 3
US-10-750-185-53947/c
; Sequence 53947, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-2
;  CURRENT APPLICATION NUMBER: US/10/750,185
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 53947
;   LENGTH: 1036
;   TYPE: DNA
;   ORGANISM: Bovine  19866880771096
US-10-750-185-53947

  Query Match             50.3%;  Score 19.6;  DB 7;  Length 1036;
  Best Local Similarity   73.5%;  Pred. No. 11;
  Matches   25;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          5 ACTCACACGCGTCTTATAGCGAGCGTCTAATTCT 38
              ||| |||| |  |||||||   || |||||||
Db        535 ACTTACACACCACTTATAGGAGGCATCTAATTTT 502


RESULT 4
US-10-750-623-53947/c
; Sequence 53947, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-1
;  CURRENT APPLICATION NUMBER: US/10/750,623
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 53947
;   LENGTH: 1036
;   TYPE: DNA
;   ORGANISM: Bovine  19866880771096
US-10-750-623-53947

  Query Match             50.3%;  Score 19.6;  DB 7;  Length 1036;
  Best Local Similarity   73.5%;  Pred. No. 11;
  Matches   25;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          5 ACTCACACGCGTCTTATAGCGAGCGTCTAATTCT 38
              ||| |||| |  |||||||   || |||||||
Db        535 ACTTACACACCACTTATAGGAGGCATCTAATTTT 502


RESULT 5
US-10-750-185-26886/c
; Sequence 26886, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-2
;  CURRENT APPLICATION NUMBER: US/10/750,185
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 26886
;   LENGTH: 2424
;   TYPE: DNA
;   ORGANISM: Bovine  19866880518817
US-10-750-185-26886

  Query Match             49.7%;  Score 19.4;  DB 7;  Length 2424;
  Best Local Similarity   70.3%;  Pred. No. 16;
  Matches   26;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          1 GAACACTCACACGCGTCTTATAGCGAGCGTCTAATTC 37
              | ||||||||||| | | | ||    ||||| | ||||
Db        621 GCACACTCACACTCATTTGATGAGCAGCGTTTTATTC 585


RESULT 6
US-10-750-623-26886/c
; Sequence 26886, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-1
;  CURRENT APPLICATION NUMBER: US/10/750,623
;  CURRENT FILING DATE:  2003-12-31
```

```
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 26886
;   LENGTH: 2424
;   TYPE: DNA
;   ORGANISM: Bovine  19866880518817
US-10-750-623-26886

  Query Match             49.7%;  Score 19.4;  DB 7;  Length 2424;
  Best Local Similarity   70.3%;  Pred. No. 16;
  Matches   26;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          1 GAACACTCACACGCGTCTTATAGCGAGCGTCTAATTC 37
              | |||||||||| | | | ||    ||||| | ||||
Db        621 GCACACTCACACTCATTTGATGAGCAGCGTTTTATTC 585


RESULT 7
US-11-136-527-485/c
; Sequence 485, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
;  APPLICANT: Wyeth
;  APPLICANT:  Mounts, William M
;  TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
;  FILE REFERENCE: 031896-041000 (AM101086)
;  CURRENT APPLICATION NUMBER: US/11/136,527
;  CURRENT FILING DATE:  2005-05-25
;  PRIOR APPLICATION NUMBER: US 60/574,294
;  PRIOR FILING DATE: 2005-05-26
;  NUMBER OF SEQ ID NOS: 362830
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 485
;   LENGTH: 3518
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
US-11-136-527-485

  Query Match             49.7%;  Score 19.4;  DB 8;  Length 3518;
  Best Local Similarity   70.3%;  Pred. No. 17;
  Matches   26;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          3 ACACTCACACGCGTCTTATAGCGAGCGTCTAATTCTA 39
              || |||||||| | | |||||||  |  || ||  ||
Db        786 ACCCTCACACGTGCCCTATAGCGCACCCCTGATGTTA 750


RESULT 8
US-10-750-185-53265
; Sequence 53265, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-2
;  CURRENT APPLICATION NUMBER: US/10/750,185
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 53265
;   LENGTH: 732
;   TYPE: DNA
;   ORGANISM: Bovine  19866881011834
US-10-750-185-53265

  Query Match             48.7%;  Score 19;  DB 7;  Length 732;
  Best Local Similarity   71.4%;  Pred. No. 20;
  Matches   25;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          2 AACACTCACACGCGTCTTATAGCGAGCGTCTAATT 36
              ||||||||  | | | ||||    ||| |||||||
Db        259 AACACTCATTCACATATTATTTTAAGCTTCTAATT 293


RESULT 9
US-10-750-623-53265
; Sequence 53265, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-1
;  CURRENT APPLICATION NUMBER: US/10/750,623
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 53265
;   LENGTH: 732
;   TYPE: DNA
;   ORGANISM: Bovine  19866881011834
US-10-750-623-53265

  Query Match             48.7%;  Score 19;  DB 7;  Length 732;
  Best Local Similarity   71.4%;  Pred. No. 20;
  Matches   25;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          2 AACACTCACACGCGTCTTATAGCGAGCGTCTAATT 36
              ||||||||  | | | ||||    ||| |||||||
Db        259 AACACTCATTCACATATTATTTTAAGCTTCTAATT 293


RESULT 10
US-10-995-561-27542
; Sequence 27542, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 27542
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-995-561-27542

  Query Match             48.2%;  Score 18.8;  DB 7;  Length 201;
  Best Local Similarity   68.4%;  Pred. No. 20;
  Matches   26;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          2 AACACTCACACGCGTCTTATAGCGAGCGTCTAATTCTA 39
              ||| ||| || || ||    |  |||| || |||||||
Db         29 AACCCTCCCATGCCTCCAGAACTGAGCCTCCAATTCTA 66
```

```
RESULT 11
US-10-995-561-27569
; Sequence 27569, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 27569
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-995-561-27569

  Query Match             48.2%;  Score 18.8;  DB 7;  Length 201;
  Best Local Similarity   68.4%;  Pred. No. 20;
  Matches   26;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          2 AACACTCACACGCGTCTTATAGCGAGCGTCTAATTCTA 39
              ||| ||| || || ||    |  |||| || |||||||
Db         18 AACCCTCCCATGCCTCCAGAACTGAGCCTCCAATTCTA 55


RESULT 12
US-10-793-626-1175
; Sequence 1175, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
;  APPLICANT: KIMMERLY, WILLIAM JOHN
;  TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
;  FILE REFERENCE: PU3480US
;  CURRENT APPLICATION NUMBER: US/10/793,626
;  CURRENT FILING DATE:  2004-03-04
;  PRIOR APPLICATION NUMBER: 60/164,258
;  PRIOR FILING DATE: 1999-11-09
;  NUMBER OF SEQ ID NOS: 4472
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1175
;   LENGTH: 537
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: synthetic
;   OTHER INFORMATION: nucleic acid sequence
US-10-793-626-1175

  Query Match             48.2%;  Score 18.8;  DB 7;  Length 537;
  Best Local Similarity   76.7%;  Pred. No. 23;
  Matches   23;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          9 ACACGCGTCTTATAGCGAGCGTCTAATTCT 38
              | |||| | ||  |||| ||||| ||||||
Db        375 ATACGCTTGTTTAAGCGTGCGTCCAATTCT 404


RESULT 13
US-10-750-185-42748
; Sequence 42748, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-2
;  CURRENT APPLICATION NUMBER: US/10/750,185
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 42748
;   LENGTH: 648
;   TYPE: DNA
;   ORGANISM: Bovine  19866881287324
US-10-750-185-42748

  Query Match             48.2%;  Score 18.8;  DB 7;  Length 648;
  Best Local Similarity   68.4%;  Pred. No. 24;
  Matches   26;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          1 GAACACTCACACGCGTCTTATAGCGAGCGTCTAATTCT 38
              ||||| | | | | ||||   |  ||||||||||  ||
Db        246 GAACAATAAAAGGGGTCTGCGAATGAGCGTCTAAAGCT 283


RESULT 14
US-10-750-623-42748
; Sequence 42748, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-1
;  CURRENT APPLICATION NUMBER: US/10/750,623
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 42748
;   LENGTH: 648
;   TYPE: DNA
;   ORGANISM: Bovine  19866881287324
US-10-750-623-42748

  Query Match             48.2%;  Score 18.8;  DB 7;  Length 648;
  Best Local Similarity   68.4%;  Pred. No. 24;
  Matches   26;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          1 GAACACTCACACGCGTCTTATAGCGAGCGTCTAATTCT 38
              ||||| | | | | ||||   |  ||||||||||  ||
Db        246 GAACAATAAAAGGGGTCTGCGAATGAGCGTCTAAAGCT 283


RESULT 15
US-10-793-626-1177/c
; Sequence 1177, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
;  APPLICANT: KIMMERLY, WILLIAM JOHN
;  TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
;  FILE REFERENCE: PU3480US
;  CURRENT APPLICATION NUMBER: US/10/793,626
;  CURRENT FILING DATE:  2004-03-04
;  PRIOR APPLICATION NUMBER: 60/164,258
;  PRIOR FILING DATE: 1999-11-09
;  NUMBER OF SEQ ID NOS: 4472
```

```
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1177
;   LENGTH: 2226
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: synthetic
;   OTHER INFORMATION: nucleic acid sequence
US-10-793-626-1177

  Query Match             48.2%;  Score 18.8;  DB 7;  Length 2226;
  Best Local Similarity   76.7%;  Pred. No. 30;
  Matches   23;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          9 ACACGCGTCTTATAGCGAGCGTCTAATTCT 38
              | |||| | ||  |||| ||||| ||||||
Db        996 ATACGCTTGTTTAAGCGTGCGTCCAATTCT 967


Search completed: January 30, 2006, 05:36:32
Job time : 392.5 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: January 29, 2006, 22:23:56 ; Search time 219 Seconds
(without alignments)
1186.863 Million cell updates/sec

Title: US-10-663-033-3-REVCOMP-4
Perfect score: 39
Sequence: 1 gaacactcacacgcgtcttatagcgagcgtctaattcta 39

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 0.1

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_21:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	22	56.4	934	2	AAX14231	Aax14231 H. pylori
	2	22	56.4	1119	3	AAA09495	Aaa09495 H. pylori
	3	21.8	55.9	2295	8	ABT18994	Abt18994 Aspergill
	4	21.8	55.9	2450	8	ABT18400	Abt18400 Aspergill
	5	21.8	55.9	2505	8	ABT20216	Abt20216 Aspergill
	6	21.8	55.9	2505	8	ABT20814	Abt20814 Aspergill
	7	21.8	55.9	4450	8	ABT17806	Abt17806 Aspergill
	8	21.8	55.9	4505	8	ABT19620	Abt19620 Aspergill
c	9	20.2	51.8	2454	8	ACA48494	Aca48494 Prokaryot
	10	20.2	51.8	29973	4	AAS29814	Aas29814 Human cyt
c	11	20.2	51.8	29973	12	ADJ12408	Adj12408 DNA fragm
c	12	20	51.3	20	3	AAA09497	Aaa09497 Primer GW
c	13	20	51.3	117754	11	ACN43866	Acn43866 Human gen
c	14	19.8	50.8	2349	13	ADT05658	Adt05658 Haemophil
	15	19.8	50.8	3469	13	ADT05413	Adt05413 Haemophil
	16	19.8	50.8	191996	13	ADT05647	Adt05647 Haemophil
c	17	19.8	50.8	263852	13	ADS99460	Ads99460 Murine ki
c	18	19.6	50.3	189158	11	ACN44124	Acn44124 Mouse gen
	19	19.4	49.7	653	2	AAV52043	Aav52043 Helicobac
	20	19.4	49.7	685	5	AAH81594	Aah81594 Human dif
c	21	19.4	49.7	2450	11	ADW22220	Adw22220 Rat hepat
c	22	19.4	49.7	3576	10	ADB58911	Adb58911 Toxicity-
c	23	19.4	49.7	3576	10	ADB53633	Adb53633 Primary r
c	24	19.4	49.7	3576	13	ADV41650	Adv41650 Rat cardi
c	25	19.4	49.7	5758	6	ABL35076	Abl35076 Murine cD
	26	19.4	49.7	35658	9	ADA02867	Ada02867 Mouse Sel
	27	19.4	49.7	35658	10	ADB72605	Adb72605 Mouse Sel
	28	19.4	49.7	35658	10	ADC85346	Adc85346 Human Itp
	29	19.4	49.7	35658	12	ADM74462	Adm74462 Murine ca
	30	19.4	49.7	38630	14	ADZ12686	Adz12686 Murine ca
c	31	19.2	49.2	505	10	ADB50969	Adb50969 Primary r
c	32	19.2	49.2	667	13	ADS61083	Ads61083 Bacterial
	33	19.2	49.2	2000	6	ABZ17024	Abz17024 Arabidops
c	34	19.2	49.2	2664	11	ACH98752	Ach98752 Klebsiell
	35	19	48.7	19	3	AAA09496	Aaa09496 Primer GW
	36	19	48.7	2363	4	ABL23910	Abl23910 Drosophil
	37	19	48.7	2670	11	ACH94275	Ach94275 Klebsiell
	38	19	48.7	2773	4	ABL04300	Abl04300 Drosophil
	39	19	48.7	3319	4	ABL04930	Abl04930 Drosophil
	40	19	48.7	3633	4	ABL04926	Abl04926 Drosophil
c	41	19	48.7	110000	14	AEB42401_28	Continuation (29 o
c	42	19	48.7	124884	5	AAH74201	Aah74201 Nucleotid
c	43	19	48.7	124884	10	ADA14878	Ada14878 Human her
c	44	19	48.7	124884	11	ADL99489	Adl99489 Varicella
c	45	19	48.7	125157	5	AAH74202	Aah74202 Nucleotid

ALIGNMENTS

RESULT 1
AAX14231/c
ID AAX14231 standard; DNA; 934 BP.
XX
AC AAX14231;
XX
DT 31-MAR-1999 (first entry)
XX
DE H. pylori GHPO 1270 gene.
XX
KW GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;
KW peptic ulcer disease; ss.
XX
OS Helicobacter pylori.
XX
FH Key Location/Qualifiers
FT CDS 51. .884
FT /*tag= a
XX
PN WO9843478-A1.
XX
PD 08-OCT-1998.
XX
PF 01-APR-1998; 98WO-US006371.
XX
PR 01-APR-1997; 97US-00833457.
PR 24-JUN-1997; 97US-00881227.
PR 29-JUL-1997; 97US-00902615.
XX
PA (INMR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Kleanthous H, Al-Garawi A, Miller C, Tomb J, Oomen RP;
XX
DR WPI; 1998-542293/46.
DR P-PSDB; AAW98512.
XX
PT New isolated Helicobacter polynucleotides - used to develop products for
PT the diagnosis, prevention and treatment of Helicobacter infections and
PT gastrointestinal diseases.
XX
PS Claim 1; Page 964-966; 2054pp; English.

```
XX
CC   This sequence represents a polynucleotide of the invention. It was
CC   isolated from Helicobacter pylori and encodes a H.pylori GHPO protein.
CC   The polypeptides can be used for preventing or treating Helicobacter
CC   infections, and gastroduodenal diseases associated with these infections,
CC   including acute, chronic, and atrophic gastritis, and peptic ulcer
CC   diseases, e.g. gastric and duodenal ulcers. They can also be used for the
CC   production of antibodies. The products can also be used for detection and
CC   diagnosis
XX
SQ   Sequence 934 BP; 306 A; 182 C; 181 G; 265 T; 0 U; 0 Other;

  Query Match             56.4%;  Score 22;  DB 2;  Length 934;
  Best Local Similarity   100.0%;  Pred. No. 7;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         18 TTATAGCGAGCGTCTAATTCTA 39
              ||||||||||||||||||||||
Db        844 TTATAGCGAGCGTCTAATTCTA 823


RESULT 2
AAA09495
ID   AAA09495 standard; DNA; 1119 BP.
XX
AC   AAA09495;
XX
DT   29-AUG-2000  (first entry)
XX
DE   H. pylori UA802 fucT2 coding sequence.
XX
KW   fucT2; alpha-1,2-fucosyltransferase; Lewis Y; glycoconjugate;
KW   fucosylated oligosaccharide; tumor associated carbohydrate antigen;
KW   tumour; ss.
XX
OS   Helicobacter pylori.
XX
FH   Key             Location/Qualifiers
FT   CDS             137. .1039
FT                   /*tag=  a
XX
PN   WO200026383-A1.
XX
PD   11-MAY-2000.
XX
PF   03-NOV-1999;   99WO-CA001031.
XX
PR   04-NOV-1998;   98US-0107268P.
PR   02-NOV-1999;   99US-00433598.
XX
PA   (UYAL-) UNIV ALBERTA.
XX
PI   Taylor DE,  Wang G,  Palcic M;
XX
DR   WPI; 2000-365628/31.
DR   P-PSDB; AAY92713.
XX
PT   Helicobacter pylori alpha1,2-fucosyltransferase enzymes useful for
PT   producing a fucosylated oligosaccharide and for diagnosing malignancies
PT   related to H. pylori infections.
XX
PS   Claim 10; Fig 1D; 71pp; English.
XX
CC   Helicobacter pylori fucT2 gene encodes alpha-1,2-fucosyltransferase. The
CC   UA802 fucT2 contains a single open reading frame encoding a protein of
CC   300 amino acids (see AAY92713). It was designated Hp fucT2 to distinguish
CC   it from the previously identified alpha-1,3-fucT, which was designated
CC   fucT. Hp fucT2 gene has a unique feature in its centre region. In
CC   addition to a poly C tract, a sequence of TAA repeats (imperfect)
CC   immediately following the poly C sequence were identified. The changes of
CC   the repeat number of both tracts contribute to the variation of the fucT2
CC   genotype (on or off status) in different strains. The UA802 fucT2 gene is
CC   intact, whereas in strain 26695 it is frameshifted. However, the 26695
CC   strain does produce a full-length protein - it was discovered that the
CC   gene contains a slippery heptanucleotide sequence identical to that found
CC   in the mRNA of Escherichia coli dnaX. This allows for ribosome slippage
CC   in translation. The Hp fucT2 polypeptide catalyzes the synthesis of Lewis
CC   Y (claimed), and other fucosylated oligosaccharides such as Lewis X,
CC   Lewis B and H type 1. The alpha-1,2-fucosyltransferase enzymes are useful
CC   for producing fucosylated oligosaccharides such as Lewis X, Lewis Y,
CC   Lewis B and H type 1, which are structurally similar to certain tumor
CC   associated carbohydrate antigens found in mammals. These glycoconjugates
CC   have research and diagnostic utility for the development of assays and
CC   reagents (e.g. antibodies) for detecting H. pylori and associated
CC   mammalian tumours
XX
SQ   Sequence 1119 BP; 363 A; 195 C; 227 G; 334 T; 0 U; 0 Other;

  Query Match             56.4%;  Score 22;  DB 3;  Length 1119;
  Best Local Similarity   100.0%;  Pred. No. 7.3;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         18 TTATAGCGAGCGTCTAATTCTA 39
              ||||||||||||||||||||||
Db       1098 TTATAGCGAGCGTCTAATTCTA 1119


RESULT 3
ABT18994
ID   ABT18994 standard; DNA; 2295 BP.
XX
AC   ABT18994;
XX
DT   16-APR-2003  (first entry)
XX
DE   Aspergillus fumigatus essential gene #1352.
XX
KW   Fungicide; cytostatic; essential gene; Aspergillus fumigatus; infection;
KW   cancer; contamination; biofilm; antibody; immune response; ds.
XX
OS   Aspergillus fumigatus.
XX
PN   WO200286090-A2.
XX
PD   31-OCT-2002.
XX
PF   23-APR-2002; 2002WO-US013142.
XX
PR   23-APR-2001; 2001US-0285697P.
PR   27-APR-2001; 2001US-0287066P.
PR   05-JUN-2001; 2001US-0295890P.
PR   09-JUL-2001; 2001US-0303899P.
PR   31-AUG-2001; 2001US-0316362P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Jiang B,  Tishkoff D,  Zamudio C,  Eroshkin AM,  Hu W,  Lemieux SM;
XX
DR   WPI; 2003-093124/08.
XX
PT   New purified or isolated nucleic acids of essential genes of Aspergillus
PT   fumigatus, useful for treating or preventing infections by A. fumigatus,
PT   or for treating a non-infectious disease in a subject e.g. cancer.
XX
PS   Disclosure; Page; 175pp; English.
XX
CC   The invention relates to novel purified or isolated nucleic acids of
CC   essential genes of Aspergillus fumigatus. The isolated nucleic acids of
CC   the invention are used to treat or prevent infections by a pathogenic
CC   organism such as A. fumigatus, to treat a non-infectious disease in a
CC   subject (e.g. cancer), to prevent or contain contamination of an object
CC   by A. fumigatus, or to prevent or inhibit formation on a surface of a
CC   biofilm comprising A. fumigatus. The polynucleotides are useful for
CC   expressing recombinant protein for characterisation, screening or
CC   therapeutic use, as markers for host tissues in which the pathogenic
CC   organisms invade or reside, for comparing with the DNA sequence of A.
```

CC fumigatus to identify duplicated genes or paralogues having the same or
CC similar biochemical activity and/or function, for comparing with DNA
CC sequences of other related or distant pathogenic organisms to identify
CC potential orthologous essential or virulence genes, for selecting and
CC making oligomers for attachment to a nucleic acid array for examination
CC of expression patterns, for raising anti-protein antibodies, as an
CC antigen to raise anti-DNA antibodies or to elicit another immune
CC response, and for identifying polynucleotides encoding the other protein
CC with which binding occurs or to identify inhibitors of the binding
CC interaction. The polypeptides may be used to raise antibodies or to
CC elicit immune response, as a reagent in assays designed to quantitatively
CC determine levels of the protein in biological fluids, as a marker for
CC host tissues in which pathogenic organism invade or reside, and to
CC isolate correlative receptors or ligands in the case or virulence
CC factors. This polynucleotide sequence represents one of the essential
CC genes of Aspergillus fumigatus of the invention
XX
SQ Sequence 2295 BP; 604 A; 607 C; 603 G; 481 T; 0 U; 0 Other;

Query Match 55.9%; Score 21.8; DB 8; Length 2295;
Best Local Similarity 78.8%; Pred. No. 10;
Matches 26; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 3 ACACTCACACGCGTCTTATAGCGAGCGTCTAAT 35
|||| || |||||||||| |||||| ||||
Db 1546 ACACGCAAACGCGTCTTAGAGCGAGGAGATAAT 1578

RESULT 4
ABT18400
ID ABT18400 standard; DNA; 2450 BP.
XX
AC ABT18400;
XX
DT 16-APR-2003 (first entry)
XX
DE Aspergillus fumigatus essential gene #758.
XX
KW Fungicide; cytostatic; essential gene; Aspergillus fumigatus; infection;
KW cancer; contamination; biofilm; antibody; immune response; ds.
XX
OS Aspergillus fumigatus.
XX
PN WO200286090-A2.
XX
PD 31-OCT-2002.
XX
PF 23-APR-2002; 2002WO-US013142.
XX
PR 23-APR-2001; 2001US-0285697P.
PR 27-APR-2001; 2001US-0287066P.
PR 05-JUN-2001; 2001US-0295890P.
PR 09-JUL-2001; 2001US-0303899P.
PR 31-AUG-2001; 2001US-0316362P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Jiang B, Tishkoff D, Zamudio C, Eroshkin AM, Hu W, Lemieux SM;
XX
DR WPI; 2003-093124/08.
XX
PT New purified or isolated nucleic acids of essential genes of Aspergillus
PT fumigatus, useful for treating or preventing infections by A. fumigatus,
PT or for treating a non-infectious disease in a subject e.g. cancer.
XX
PS Disclosure; Page; 175pp; English.
XX
CC The invention relates to novel purified or isolated nucleic acids of
CC essential genes of Aspergillus fumigatus. The isolated nucleic acids of
CC the invention are used to treat or prevent infections by a pathogenic
CC organism such as A. fumigatus, to treat a non-infectious disease in a
CC subject (e.g. cancer), to prevent or contain contamination of an object
CC by A. fumigatus, or to prevent or inhibit formation on a surface of a
CC biofilm comprising A. fumigatus. The polynucleotides are useful for
CC expressing recombinant protein for characterisation, screening or
CC therapeutic use, as markers for host tissues in which the pathogenic
CC organisms invade or reside, for comparing with the DNA sequence of A.
CC fumigatus to identify duplicated genes or paralogues having the same or
CC similar biochemical activity and/or function, for comparing with DNA
CC sequences of other related or distant pathogenic organisms to identify
CC potential orthologous essential or virulence genes, for selecting and
CC making oligomers for attachment to a nucleic acid array for examination
CC of expression patterns, for raising anti-protein antibodies, as an
CC antigen to raise anti-DNA antibodies or to elicit another immune
CC response, and for identifying polynucleotides encoding the other protein
CC with which binding occurs or to identify inhibitors of the binding
CC interaction. The polypeptides may be used to raise antibodies or to
CC elicit immune response, as a reagent in assays designed to quantitatively
CC determine levels of the protein in biological fluids, as a marker for
CC host tissues in which pathogenic organism invade or reside, and to
CC isolate correlative receptors or ligands in the case or virulence
CC factors. This polynucleotide sequence represents one of the essential
CC genes of Aspergillus fumigatus of the invention
XX
SQ Sequence 2450 BP; 644 A; 673 C; 628 G; 505 T; 0 U; 0 Other;

Query Match 55.9%; Score 21.8; DB 8; Length 2450;
Best Local Similarity 78.8%; Pred. No. 10;
Matches 26; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 3 ACACTCACACGCGTCTTATAGCGAGCGTCTAAT 35
|||| || |||||||||| |||||| ||||
Db 1546 ACACGCAAACGCGTCTTAGAGCGAGGAGATAAT 1578

RESULT 5
ABT20216
ID ABT20216 standard; DNA; 2505 BP.
XX
AC ABT20216;
XX
DT 16-APR-2003 (first entry)
XX
DE Aspergillus fumigatus essential gene #2574.
XX
KW Fungicide; cytostatic; essential gene; Aspergillus fumigatus; infection;
KW cancer; contamination; biofilm; antibody; immune response; ds.
XX
OS Aspergillus fumigatus.
XX
PN WO200286090-A2.
XX
PD 31-OCT-2002.
XX
PF 23-APR-2002; 2002WO-US013142.
XX
PR 23-APR-2001; 2001US-0285697P.
PR 27-APR-2001; 2001US-0287066P.
PR 05-JUN-2001; 2001US-0295890P.
PR 09-JUL-2001; 2001US-0303899P.
PR 31-AUG-2001; 2001US-0316362P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Jiang B, Tishkoff D, Zamudio C, Eroshkin AM, Hu W, Lemieux SM;
XX
DR WPI; 2003-093124/08.
XX
PT New purified or isolated nucleic acids of essential genes of Aspergillus
PT fumigatus, useful for treating or preventing infections by A. fumigatus,
PT or for treating a non-infectious disease in a subject e.g. cancer.
XX
PS Disclosure; Page; 175pp; English.
XX
CC The invention relates to novel purified or isolated nucleic acids of
CC essential genes of Aspergillus fumigatus. The isolated nucleic acids of

```
CC   the invention are used to treat or prevent infections by a pathogenic
CC   organism such as A. fumigatus, to treat a non-infectious disease in a
CC   subject (e.g. cancer), to prevent or contain contamination of an object
CC   by A. fumigatus, or to prevent or inhibit formation on a surface of a
CC   biofilm comprising A. fumigatus. The polynucleotides are useful for
CC   expressing recombinant protein for characterisation, screening or
CC   therapeutic use, as markers for host tissues in which the pathogenic
CC   organisms invade or reside, for comparing with the DNA sequence of A.
CC   fumigatus to identify duplicated genes or paralogues having the same or
CC   similar biochemical activity and/or function, for comparing with DNA
CC   sequences of other related or distant pathogenic organisms to identify
CC   potential orthologous essential or virulence genes, for selecting and
CC   making oligomers for attachment to a nucleic acid array for examination
CC   of expression patterns, for raising anti-protein antibodies, as an
CC   antigen to raise anti-DNA antibodies or to elicit another immune
CC   response, and for identifying polynucleotides encoding the other protein
CC   with which binding occurs or to identify inhibitors of the binding
CC   interaction. The polypeptides may be used to raise antibodies or to
CC   elicit immune response, as a reagent in assays designed to quantitatively
CC   determine levels of the protein in biological fluids, as a marker for
CC   host tissues in which pathogenic organism invade or reside, and to
CC   isolate correlative receptors or ligands in the case or virulence
CC   factors. This polynucleotide sequence represents one of the essential
CC   genes of Aspergillus fumigatus of the invention
XX
SQ   Sequence 2505 BP; 641 A; 675 C; 677 G; 512 T; 0 U; 0 Other;

  Query Match             55.9%;  Score 21.8;  DB 8;  Length 2505;
  Best Local Similarity   78.8%;  Pred. No. 10;
  Matches   26;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          3 ACACTCACACGCGTCTTATAGCGAGCGTCTAAT 35
              |||| || |||||||||| ||||||    ||||
Db       1471 ACACGCAAACGCGTCTTAGAGCGAGGAGATAAT 1503


RESULT 6
ABT20814
ID   ABT20814 standard; DNA; 2505 BP.
XX
AC   ABT20814;
XX
DT   16-APR-2003  (first entry)
XX
DE   Aspergillus fumigatus essential gene #3172.
XX
KW   Fungicide; cytostatic; essential gene; Aspergillus fumigatus; infection;
KW   cancer; contamination; biofilm; antibody; immune response; ds.
XX
OS   Aspergillus fumigatus.
XX
PN   WO200286090-A2.
XX
PD   31-OCT-2002.
XX
PF   23-APR-2002; 2002WO-US013142.
XX
PR   23-APR-2001; 2001US-0285697P.
PR   27-APR-2001; 2001US-0287066P.
PR   05-JUN-2001; 2001US-0295890P.
PR   09-JUL-2001; 2001US-0303899P.
PR   31-AUG-2001; 2001US-0316362P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Jiang B,  Tishkoff D,  Zamudio C,  Eroshkin AM,  Hu W,  Lemieux SM;
XX
DR   WPI; 2003-093124/08.
XX
PT   New purified or isolated nucleic acids of essential genes of Aspergillus
PT   fumigatus, useful for treating or preventing infections by A. fumigatus,
PT   or for treating a non-infectious disease in a subject e.g. cancer.
XX
PS   Disclosure; Page; 175pp; English.
XX
CC   The invention relates to novel purified or isolated nucleic acids of
CC   essential genes of Aspergillus fumigatus. The isolated nucleic acids of
CC   the invention are used to treat or prevent infections by a pathogenic
CC   organism such as A. fumigatus, to treat a non-infectious disease in a
CC   subject (e.g. cancer), to prevent or contain contamination of an object
CC   by A. fumigatus, or to prevent or inhibit formation on a surface of a
CC   biofilm comprising A. fumigatus. The polynucleotides are useful for
CC   expressing recombinant protein for characterisation, screening or
CC   therapeutic use, as markers for host tissues in which the pathogenic
CC   organisms invade or reside, for comparing with the DNA sequence of A.
CC   fumigatus to identify duplicated genes or paralogues having the same or
CC   similar biochemical activity and/or function, for comparing with DNA
CC   sequences of other related or distant pathogenic organisms to identify
CC   potential orthologous essential or virulence genes, for selecting and
CC   making oligomers for attachment to a nucleic acid array for examination
CC   of expression patterns, for raising anti-protein antibodies, as an
CC   antigen to raise anti-DNA antibodies or to elicit another immune
CC   response, and for identifying polynucleotides encoding the other protein
CC   with which binding occurs or to identify inhibitors of the binding
CC   interaction. The polypeptides may be used to raise antibodies or to
CC   elicit immune response, as a reagent in assays designed to quantitatively
CC   determine levels of the protein in biological fluids, as a marker for
CC   host tissues in which pathogenic organism invade or reside, and to
CC   isolate correlative receptors or ligands in the case or virulence
CC   factors. This polynucleotide sequence represents one of the essential
CC   genes of Aspergillus fumigatus of the invention
XX
SQ   Sequence 2505 BP; 641 A; 675 C; 677 G; 512 T; 0 U; 0 Other;

  Query Match             55.9%;  Score 21.8;  DB 8;  Length 2505;
  Best Local Similarity   78.8%;  Pred. No. 10;
  Matches   26;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          3 ACACTCACACGCGTCTTATAGCGAGCGTCTAAT 35
              |||| || |||||||||| ||||||    ||||
Db       1471 ACACGCAAACGCGTCTTAGAGCGAGGAGATAAT 1503


RESULT 7
ABT17806
ID   ABT17806 standard; DNA; 4450 BP.
XX
AC   ABT17806;
XX
DT   16-APR-2003  (first entry)
XX
DE   Aspergillus fumigatus essential gene #164.
XX
KW   Fungicide; cytostatic; essential gene; Aspergillus fumigatus; infection;
KW   cancer; contamination; biofilm; antibody; immune response; ds.
XX
OS   Aspergillus fumigatus.
XX
PN   WO200286090-A2.
XX
PD   31-OCT-2002.
XX
PF   23-APR-2002; 2002WO-US013142.
XX
PR   23-APR-2001; 2001US-0285697P.
PR   27-APR-2001; 2001US-0287066P.
PR   05-JUN-2001; 2001US-0295890P.
PR   09-JUL-2001; 2001US-0303899P.
PR   31-AUG-2001; 2001US-0316362P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Jiang B,  Tishkoff D,  Zamudio C,  Eroshkin AM,  Hu W,  Lemieux SM;
XX
DR   WPI; 2003-093124/08.
XX
```

PT New purified or isolated nucleic acids of essential genes of Aspergillus
PT fumigatus, useful for treating or preventing infections by A. fumigatus,
PT or for treating a non-infectious disease in a subject e.g. cancer.
XX
PS Disclosure; Page; 175pp; English.
XX
CC The invention relates to novel purified or isolated nucleic acids of
CC essential genes of Aspergillus fumigatus. The isolated nucleic acids of
CC the invention are used to treat or prevent infections by a pathogenic
CC organism such as A. fumigatus, to treat a non-infectious disease in a
CC subject (e.g. cancer), to prevent or contain contamination of an object
CC by A. fumigatus, or to prevent or inhibit formation on a surface of a
CC biofilm comprising A. fumigatus. The polynucleotides are useful for
CC expressing recombinant protein for characterisation, screening or
CC therapeutic use, as markers for host tissues in which the pathogenic
CC organisms invade or reside, for comparing with the DNA sequence of A.
CC fumigatus to identify duplicated genes or paralogues having the same or
CC similar biochemical activity and/or function, for comparing with DNA
CC sequences of other related or distant pathogenic organisms to identify
CC potential orthologous essential or virulence genes, for selecting and
CC making oligomers for attachment to a nucleic acid array for examination
CC of expression patterns, for raising anti-protein antibodies, as an
CC antigen to raise anti-DNA antibodies or to elicit another immune
CC response, and for identifying polynucleotides encoding the other protein
CC with which binding occurs or to identify inhibitors of the binding
CC interaction. The polypeptides may be used to raise antibodies or to
CC elicit immune response, as a reagent in assays designed to quantitatively
CC determine levels of the protein in biological fluids, as a marker for
CC host tissues in which pathogenic organism invade or reside, and to
CC isolate correlative receptors or ligands in the case or virulence
CC factors. This polynucleotide sequence represents one of the essential
CC genes of Aspergillus fumigatus of the invention
XX
SQ Sequence 4450 BP; 1174 A; 1126 C; 1152 G; 998 T; 0 U; 0 Other;

Query Match 55.9%; Score 21.8; DB 8; Length 4450;
Best Local Similarity 78.8%; Pred. No. 12;
Matches 26; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
Qy          3 ACACTCACACGCGTCTTATAGCGAGCGTCTAAT 35
              |||| || |||||||||| ||||||    ||||
Db       2546 ACACGCAAACGCGTCTTAGAGCGAGGAGATAAT 2578
```

RESULT 8
ABT19620
ID ABT19620 standard; DNA; 4505 BP.
XX
AC ABT19620;
XX
DT 16-APR-2003 (first entry)
XX
DE Aspergillus fumigatus essential gene #1978.
XX
KW Fungicide; cytostatic; essential gene; Aspergillus fumigatus; infection;
KW cancer; contamination; biofilm; antibody; immune response; ds.
XX
OS Aspergillus fumigatus.
XX
PN WO200286090-A2.
XX
PD 31-OCT-2002.
XX
PF 23-APR-2002; 2002WO-US013142.
XX
PR 23-APR-2001; 2001US-0285697P.
PR 27-APR-2001; 2001US-0287066P.
PR 05-JUN-2001; 2001US-0295890P.
PR 09-JUL-2001; 2001US-0303899P.
PR 31-AUG-2001; 2001US-0316362P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Jiang B, Tishkoff D, Zamudio C, Broshkin AM, Hu W, Lemieux SM;
XX
DR WPI; 2003-093124/08.
XX
PT New purified or isolated nucleic acids of essential genes of Aspergillus
PT fumigatus, useful for treating or preventing infections by A. fumigatus,
PT or for treating a non-infectious disease in a subject e.g. cancer.
XX
PS Disclosure; Page; 175pp; English.
XX
CC The invention relates to novel purified or isolated nucleic acids of
CC essential genes of Aspergillus fumigatus. The isolated nucleic acids of
CC the invention are used to treat or prevent infections by a pathogenic
CC organism such as A. fumigatus, to treat a non-infectious disease in a
CC subject (e.g. cancer), to prevent or contain contamination of an object
CC by A. fumigatus, or to prevent or inhibit formation on a surface of a
CC biofilm comprising A. fumigatus. The polynucleotides are useful for
CC expressing recombinant protein for characterisation, screening or
CC therapeutic use, as markers for host tissues in which the pathogenic
CC organisms invade or reside, for comparing with the DNA sequence of A.
CC fumigatus to identify duplicated genes or paralogues having the same or
CC similar biochemical activity and/or function, for comparing with DNA
CC sequences of other related or distant pathogenic organisms to identify
CC potential orthologous essential or virulence genes, for selecting and
CC making oligomers for attachment to a nucleic acid array for examination
CC of expression patterns, for raising anti-protein antibodies, as an
CC antigen to raise anti-DNA antibodies or to elicit another immune
CC response, and for identifying polynucleotides encoding the other protein
CC with which binding occurs or to identify inhibitors of the binding
CC interaction. The polypeptides may be used to raise antibodies or to
CC elicit immune response, as a reagent in assays designed to quantitatively
CC determine levels of the protein in biological fluids, as a marker for
CC host tissues in which pathogenic organism invade or reside, and to
CC isolate correlative receptors or ligands in the case or virulence
CC factors. This polynucleotide sequence represents one of the essential
CC genes of Aspergillus fumigatus of the invention
XX
SQ Sequence 4505 BP; 1211 A; 1128 C; 1169 G; 997 T; 0 U; 0 Other;

Query Match 55.9%; Score 21.8; DB 8; Length 4505;
Best Local Similarity 78.8%; Pred. No. 12;
Matches 26; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
Qy          3 ACACTCACACGCGTCTTATAGCGAGCGTCTAAT 35
              |||| || |||||||||| ||||||    ||||
Db       2471 ACACGCAAACGCGTCTTAGAGCGAGGAGATAAT 2503
```

RESULT 9
ACA48494/c
ID ACA48494 standard; DNA; 2454 BP.
XX
AC ACA48494;
XX
DT 19-JUN-2003 (first entry)
XX
DE Prokaryotic essential gene #30151.
XX
KW Antisense; ds; prokaryotic essential gene; cell proliferation;
KW drug design; gene.
XX
OS Streptococcus mutans.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
DR P-PSDB; ABU44624.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 14; SEQ ID NO 36364; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide;(5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2454 BP; 723 A; 471 C; 523 G; 737 T; 0 U; 0 Other;

Query Match 51.8%; Score 20.2; DB 8; Length 2454;
Best Local Similarity 88.0%; Pred. No. 57;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 AACACTCACACGCGTCTTATAGCGA 26
||||||| ||||||| |||||||
Db 1572 AACACTCATACGCGTCAAATAGCGA 1548

RESULT 10
AAS29814
ID AAS29814 standard; DNA; 29973 BP.
XX
AC AAS29814;
XX
DT 21-NOV-2001 (first entry)
XX
DE Human cytoskeletal element-related polypeptide encoding genomic DNA #1.
XX
KW Cytoskeletal element-related protein; human; mouse; rabbit; goat; horse;
KW cat; dog; chicken; sheep; immunosuppressive; antiarthritic; vasotropic;
KW antirheumatic; antiproliferative; cytostatic; cardiant; neuroprotective;
KW cerebroprotective; nootropic; antibacterial; virucide; fungicide; cancer;
KW ophthalmological; vulnerary; gene therapy; autoimmune disease; neoplasm;
KW hyperproliferative disorder; breast; liver; cardiovascular disorder; ds;
KW cerebrovascular disorder; nervous system disorder; bacterial infection;
KW fungal infection; viral infection; ocular disorder; endocrine disorder;
KW gastrointestinal disorder; renal disorder; respiratory disorder;
KW wound healing; skin aging; organ transplantation; food preservative;
KW tissue regeneration; anti-infertility; food additive.
XX
OS Homo sapiens.
XX
PN WO200155168-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001331.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.

PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-476182/51.
XX
PT Novel isolated human cytoskeletal element-related polypeptide useful for
PT diagnosis/treatment of neoplastic disorders, disorders associated with
PT neural transmission, chromosomal abnormalities, autoimmune disorders.
XX
PS Claim 1; SEQ ID NO 99; 505pp; English.
XX
CC Sequences AAS29814-AAS29838 represent genomic DNA molecules, which encode
CC the cytoskeletal element-related polypeptides of the invention.
CC Cytoskeletal polypeptides and their associated polynucleotides are useful
CC in the diagnosis, treatment and prevention of various types of disorders
CC in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or
CC sheep. A pathological condition can be determined by determining the
CC presence or absence of a mutation in a cytoskeletal polynucleotide. The
CC treatable disorders include autoimmune diseases such as rheumatoid
CC arthritis, hyperproliferative disorders such as neoplasms of the breast
CC or liver, cardiovascular disorders such as cardiac arrest,
CC cerebrovascular disorders such as cerebral ischaemia, nervous system
CC disorders such as Alzheimer's disease, infections caused by bacteria,
CC viruses and fungi, ocular disorders such as corneal infection, endocrine
CC disorders such as premature labour and infertility, gastrointestinal
CC disorders such as Crohn's disease, renal disorders such as
CC glomerulonephritis and respiratory disorders such as asthma. The
CC polypeptides can also be used to aid wound healing, to prevent skin aging
CC due to sunburn, to maintain organs before transplantation, to regenerate
CC tissues and in chemotaxis. The polypeptides can also be used as a food
CC additive or preservative to increase or decrease storage capabilities.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences

Query Match 51.8%; Score 20.2; DB 4; Length 29973;
Best Local Similarity 75.8%; Pred. No. 91;
Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```
Qy          6 CTCACACGCGTCTTATAGCGAGCGTCTAATTCT 38
              |||| || ||||| | || ||| ||||| |||
Db      25312 CTCAGCCGTGTCTTCTGGCAAGCCTCTAACTCT 25344
```

RESULT 11
ADJ12408/c
ID ADJ12408 standard; DNA; 29973 BP.
XX
AC ADJ12408;
XX
DT 20-MAY-2004 (first entry)
XX
DE DNA fragment of a BAC clone that encodes a human secreted protein Seq262.
XX
KW human; secreted; cancer; haematopoietic disease; anaemia;
KW multiple myeloma; reproductive system disorder; prostatitis;
KW inguinal hernia; musculoskeletal disease; systemic lupus erythematosus;
KW gout; cardiovascular disease; arrhythmia; hypernatraemia; fetal disease;
KW fetal alcohol syndrome; Down's syndrome; excretory disease;
KW urinary incontinence; renal disorder; neural; sensory disease;
KW Alzheimer's disease; meningitis; respiratory disease; emphysema;
KW occupational lung disease; endocrine disease; diabetes;
KW glomerulonephritis; digestive disease; portal hypertension;
KW irritable bowel syndrome; epithelial disease; scleroderma;
KW epidermolysis bullosa; cytostatic; antianemic; antiarthritic;

KW antiasthmatic; anti-HIV; immunosuppressive; antiinflammatory;
KW antipsoriatic; antibacterial; osteopathic; dermatological; antigout;
KW immunomodulator; antiarrhythmic; cardiant; nootropic; antilipemic;
KW nephrotropic; uropathic; neuroprotective; antiparkinsonian; tranquilizer;
KW antidiabetic; anabolic; hypertensive; vulnerary; ds.
XX
OS Homo sapiens.
XX
PN US2004010132-A1.
XX
PD 15-JAN-2004.
XX
PF 30-OCT-2001; 2001US-00984429.
XX
PR 09-OCT-1997; 97US-0061463P.
PR 09-OCT-1997; 97US-0061527P.
PR 09-OCT-1997; 97US-0061529P.
PR 09-OCT-1997; 97US-0061532P.
PR 09-OCT-1997; 97US-0061536P.
PR 09-OCT-1997; 97US-0071498P.
PR 08-OCT-1998; 98WO-US021142.
PR 08-APR-1999; 99US-00288143.
PR 01-NOV-2000; 2000US-0244591P.
XX
PA (ROSE/) ROSEN C A.
PA (BREW/) BREWER L A.
PA (DUAN/) DUAN R D.
PA (RUBE/) RUBEN S M.
PA (FLOR/) FLORENCE K A.
PA (GREE/) GREENE J M.
PA (YOUN/) YOUNG P E.
PA (FERR/) FERRIE A M.
PA (YUGG/) YU G.
PA (FLOR/) FLORENCE C.
PA (EBNE/) EBNER R.
PA (OLSE/) OLSEN H.
XX
PI Rosen CA, Brewer LA, Duan RD, Ruben SM, Florence KA, Greene JM;
PI Young PE, Ferrie AM, Yu G, Florence C, Ebner R, Olsen H;
XX
DR WPI; 2004-090518/09.
XX
PT New isolated nucleic acids and polypeptides, useful for diagnosing,
PT treating, preventing or ameliorating diseases or disorders e.g. cancer,
PT anemia, arthritis, asthma, inflammatory bowel disease or Alzheimer's
PT disease.
XX
PS Disclosure; SEQ ID NO 262; 286pp; English.
XX
CC This invention relates to novel polynucleotides encoding human secreted
CC proteins. Specifically, it refers to the vectors, host cells, recombinant
CC and synthetic methods for producing human polynucleotides, polypeptides
CC and antibodies. Furthermore, it relates to screening methods to identify
CC agonists and antagonists that can be used to inhibit or enhance the
CC production and function of the secreted proteins. The present invention
CC describes these compositions as useful for diagnosing, treating or
CC preventing disorders such as cancer, haematopoietic diseases including
CC anaemia and multiple myeloma, reproductive system disorders including
CC prostatitis and inguinal hernia, musculoskeletal diseases including
CC systemic lupus erythematosus and gout, cardiovascular disease including
CC arrhythmia and hypernatraemia, mixed fetal diseases including fetal
CC alcohol syndrome and Down's syndrome, excretory diseases including
CC urinary incontinence and renal disorders, neural or sensory disease
CC including Alzheimer's disease and meningitis, respiratory disease
CC including emphysema and occupational lung disease, endocrine diseases
CC including diabetes and glomerulonephritis, digestive diseases including
CC portal hypertension and irritable bowel syndrome and connective tissue or
CC epithelial diseases including scleroderma and epidermolysis bullosa. As
CC such, there are various activites such as cytostatic, antianemic,
CC antiarthritic, antiasthmatic, anti-HIV, immunosuppressive,
CC antiinflammatory, antipsoriatic, antibacterial, osteopathic,
CC dermatological, antigout, immunomodulator, antiarrhythmic, cardiant,
CC nootropic, antilipemic, nephrotropic, uropathic, neuroprotective,
CC antiparkinsonian, tranquilizer, antidiabetic, anabolic, hypertensive and
CC vulnerary. This polynucleotide is a DNA fragment of a BAC clone that
CC encodes a human secreted protein of the invention. NOTE: This sequence
CC does not appear in the printed specification but has been obtained in
CC electronic format from the US patent office at the following web site
CC www.seqdata.uspto.gov/sequence.html; Document ID: 20040010132.
XX
SQ Sequence 29973 BP; 7911 A; 6282 C; 7047 G; 8733 T; 0 U; 0 Other;

Query Match 51.8%; Score 20.2; DB 12; Length 29973;
Best Local Similarity 75.8%; Pred. No. 91;
Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 6 CTCACACGCGTCTTATAGCGAGCGTCTAATTCT 38
|||| || ||||| | || ||| ||||| |||
Db 4662 CTCAGCCGTGTCTTCTGGCAAGCCTCTAACTCT 4630

RESULT 12
AAA09497/c
ID AAA09497 standard; DNA; 20 BP.
XX
AC AAA09497;
XX
DT 29-AUG-2000 (first entry)
XX
DE Primer GW32 for Hp fucT2 gene amplification.
XX
KW Primer; fucT2; alpha-1,2-fucosyltransferase; Lewis Y; glycoconjugate;
KW fucosylated oligosaccharide; tumor associated carbohydrate antigen;
KW tumour; ss.
XX
OS Helicobacter pylori.
XX
PN WO200026383-A1.
XX
PD 11-MAY-2000.
XX
PF 03-NOV-1999; 99WO-CA001031.
XX
PR 04-NOV-1998; 98US-0107268P.
PR 02-NOV-1999; 99US-00433598.
XX
PA (UYAL-) UNIV ALBERTA.
XX
PI Taylor DE, Wang G, Palcic M;
XX
DR WPI; 2000-365628/31.
XX
PT Helicobacter pylori alpha1,2-fucosyltransferase enzymes useful for
PT producing a fucosylated oligosaccharide and for diagnosing malignancies
PT related to H. pylori infections.
XX
PS Example 1; Page 30; 71pp; English.
XX
CC AAA09496-97 are primers corresponding to positions 99980-99962 and 98855-
CC 98874, respectively, in the published Helicobacter pylori genome. They
CC were used to PCR amplify a 1.12 kb fragment containing the H. pylori
CC fucT2 coding sequence (encoding alpha-1,2-fucosyltransferase) from
CC strains 26695 and UA802. The UA802 fucT2 (see AAA09495) contains a single
CC open reading frame encoding a protein of 300 amino acids (see AAY92713).
CC It was designated Hp fucT2 to distinguish it from the previously
CC identified alpha-1,3-fucT, which was designated fucT. Hp fucT2 gene has a
CC unique feature in its centre region. In addition to a poly C tract, a
CC sequence of TAA repeats (imperfect) immediately following the poly C
CC sequence were identified. The changes of the repeat number of both tracts
CC contribute to the variation of the fucT2 genotype (on or off status) in
CC different strains. The Hp fucT2 polypeptide catalyzes the synthesis of
CC Lewis Y (claimed), and other fucosylated oligosaccharides such as Lewis
CC X, Lewis B and H type 1. The alpha-1,2-fucosyltransferase enzymes are
CC useful for producing fucosylated oligosaccharides such as Lewis X, Lewis
CC Y, Lewis B and H type 1, which are structurally similar to certain tumor
CC associated carbohydrate antigens found in mammals. These glycoconjugates

```
CC   have research and diagnostic utility for the development of assays and
CC   reagents (e.g. antibodies) for detecting H. pylori and associated
CC   mammalian tumours
XX
SQ   Sequence 20 BP; 6 A; 4 C; 4 G; 6 T; 0 U; 0 Other;

  Query Match             51.3%;  Score 20;  DB 3;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 29;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         20 ATAGCGAGCGTCTAATTCTA 39
              ||||||||||||||||||||
Db         20 ATAGCGAGCGTCTAATTCTA 1


RESULT 13
ACN43866/c
ID   ACN43866 standard; DNA; 117754 BP.
XX
AC   ACN43866;
XX
DT   18-NOV-2004  (first entry)
XX
DE   Human genomic sequence hCG37475.
XX
KW   Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO2003073826-A2.
XX
PD   12-SEP-2003.
XX
PF   28-FEB-2003; 2003WO-US006235.
XX
PR   01-MAR-2002; 2002US-00087192.
XX
PA   (SAGR-) SAGRES DISCOVERY.
XX
PI   Morris DW;
XX
DR   WPI; 2003-328604/31.
XX
PT   Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
PT   comprises a nucleotide sequence.
XX
PS   Claim 1; SEQ ID NO 28; 0pp; English.
XX
CC   The present invention relates to novel DNA and protein sequences which
CC   are associated with carcinomas. The sequences are useful for: (i) for
CC   screening drug candidates; (ii) for screening of bioactive agent capable
CC   of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
CC   a bioactive agent capable of modulating the activity of CAP; (iv) for
CC   evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
CC   carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
CC   carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
CC   (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
CC   determining Carcinoma Associated (CA) gene copy number. In addition, the
CC   CA genes are useful as DNA vaccines and the CAP are useful as markers of
CC   carcinoma including lymphoma. The present sequence is one such CA coding
CC   sequence. Note: This patent is an equivalent to basic patent
CC   US2002182586A1, for which no sequence data was published
XX
SQ   Sequence 117754 BP; 36841 A; 19567 C; 21519 G; 39647 T; 0 U; 180 Other;

  Query Match             51.3%;  Score 20;  DB 11;  Length 117754;
  Best Local Similarity   72.2%;  Pred. No. 1.5e+02;
  Matches   26;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          3 ACACTCACACGCGTCTTATAGCGAGCGTCTAATTCT 38
              |||| ||||| | ||||| || ||    || |||||
Db      35011 ACACACACACACATCTTAGAGGGATTTGCTCATTCT 34976


RESULT 14
ADT05658/c
ID   ADT05658 standard; DNA; 2349 BP.
XX
AC   ADT05658;
XX
DT   02-DEC-2004  (first entry)
XX
DE   Haemophilus influenzae (NTHi) protein coding sequence - SEQ ID 694.
XX
KW   middle ear bacterial infection; nasopharynx bacterial infection; gene;
KW   ds.
XX
OS   Haemophilus influenzae.
XX
PN   WO2004078949-A2.
XX
PD   16-SEP-2004.
XX
PF   05-MAR-2004; 2004WO-US007001.
XX
PR   06-MAR-2003; 2003US-0453134P.
XX
PA   (CHIL-) CHILDRENS HOSPITAL INC.
XX
PI   Bakaletz LO,  Munson RS,  Dyer DW;
XX
DR   WPI; 2004-662422/64.
DR   P-PSDB; ADT05659.
XX
PT   New polynucleotides of nontypeable strain of Haemophilus influenzae,
PT   useful for treating or preventing NTHi bacterial infections of the middle
PT   ear and/or nasopharynx.
XX
PS   Claim 1; SEQ ID NO 694; 88pp; English.
XX
CC   The invention comprises nucleotide sequences (genes) from the genome of a
CC   nontypeable strain of Haemophilus influenzae (NTHi). The NTHi DNA
CC   sequences of the invention are useful for treating or preventing NTHi
CC   bacterial infections of the middle ear and/or nasopharynx. The present
CC   DNA sequence represents an NTHi protein coding sequence of the invention.
XX
SQ   Sequence 2349 BP; 741 A; 431 C; 527 G; 650 T; 0 U; 0 Other;

  Query Match             50.8%;  Score 19.8;  DB 13;  Length 2349;
  Best Local Similarity   77.4%;  Pred. No. 87;
  Matches   24;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          8 CACACGCGTCTTATAGCGAGCGTCTAATTCT 38
              |||| ||||   ||| |||||| |||||||
Db       1233 CACAAGCGTAGAATAACGAGCGCGTAATTCT 1203


RESULT 15
ADT05413
ID   ADT05413 standard; DNA; 3469 BP.
XX
AC   ADT05413;
XX
DT   02-DEC-2004  (first entry)
XX
DE   Haemophilus influenzae (NTHi) contig DNA sequence - SEQ ID 449.
XX
KW   middle ear bacterial infection; nasopharynx bacterial infection; ds;
KW   contig.
XX
OS   Haemophilus influenzae.
XX
PN   WO2004078949-A2.
XX
PD   16-SEP-2004.
XX
```

PF 05-MAR-2004; 2004WO-US007001.
XX
PR 06-MAR-2003; 2003US-0453134P.
XX
PA (CHIL-) CHILDRENS HOSPITAL INC.
XX
PI Bakaletz LO, Munson RS, Dyer DW;
XX
DR WPI; 2004-662422/64.
XX
PT New polynucleotides of nontypeable strain of Haemophilus influenzae,
PT useful for treating or preventing NTHi bacterial infections of the middle
PT ear and/or nasopharynx.
XX
PS Example 1; SEQ ID NO 449; 88pp; English.
XX
CC The invention comprises nucleotide sequences (genes) from the genome of a
CC nontypeable strain of Haemophilus influenzae (NTHi). The NTHi DNA
CC sequences of the invention are useful for treating or preventing NTHi
CC bacterial infections of the middle ear and/or nasopharynx. The present
CC nucleic acid represents an NTHi contig sequence of the invention.
XX
SQ Sequence 3469 BP; 962 A; 761 C; 667 G; 1079 T; 0 U; 0 Other;

Query Match 50.8%; Score 19.8; DB 13; Length 3469;
Best Local Similarity 77.4%; Pred. No. 93;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
Qy          8 CACACGCGTCTTATAGCGAGCGTCTAATTCT 38
              |||| ||||   ||| ||||||  |||||||
Db       2149 CACAAGCGTAGAATAACGAGCGCGTAATTCT 2179
```

Search completed: January 30, 2006, 02:04:23
Job time : 222 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: January 29, 2006, 22:30:51 ; Search time 894 Seconds
(without alignments)
2479.748 Million cell updates/sec

Title: US-10-663-033-REVCOMP4-SEQ3
Perfect score: 39
Sequence: 1 atagcgagcgtctaattctagaacactcacacgcgtctt 39

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 0.1

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	24	61.5	5145	1	AF364060	AF364060 Cyanothec
	2	22.8	58.5	281517	14	AC127144	AC127144 Rattus no
	3	22.4	57.4	653	6	BD010597	BD010597 Helicobac
	4	22.4	57.4	1284	15	AK100414	AK100414 Oryza sat
	5	22.4	57.4	3491	15	AK065982	AK065982 Oryza sat
c	6	22.4	57.4	10530	1	AE000531	AE000531 Helicobac
	7	22.4	57.4	110000	15	AP008210_304	Continuation (305
	8	22.4	57.4	166804	15	OSJN00077	AL606638 Oryza sat
c	9	22.2	56.9	168482	9	AC145082	AC145082 Mus muscu
c	10	22.2	56.9	196824	9	AC096623	AC096623 Mus muscu
	11	21.8	55.9	613	10	BV317977	BV317977 S236P6501
c	12	21.8	55.9	84669	14	AC153673	AC153673 Bos tauru
	13	21.8	55.9	168735	14	AP001852	AP001852 Homo sapi
c	14	21.8	55.9	177575	14	AC021857	AC021857 Homo sapi
c	15	21.8	55.9	201519	8	AC099689	AC099689 Homo sapi
c	16	21.6	55.4	301761	1	AE016945	AE016945 Bacteroid
	17	21.4	54.9	444	15	U72384	U72384 Azorella tr
	18	21.4	54.9	2182	6	AR508487	AR508487 Sequence
c	19	21.4	54.9	282611	14	AC156719	AC156719 Bos tauru
	20	21.2	54.4	612	10	BV043572	BV043572 S209P6000
c	21	21.2	54.4	11287	9	AF034692	AF034692 Mus muscu
c	22	21.2	54.4	43501	8	AC005783	AC005783 Homo sapi
	23	21.2	54.4	143428	8	AL162587	AL162587 Human DNA
	24	21.2	54.4	162062	9	AL691496	AL691496 Mouse DNA
c	25	21.2	54.4	201997	14	AC141220	AC141220 Rattus no
c	26	21.2	54.4	237997	14	AC093989	AC093989 Rattus no
	27	21	53.8	71145	14	AC016088	AC016088 Homo sapi
c	28	21	53.8	105616	9	AL935134	AL935134 Mouse DNA
c	29	21	53.8	107458	8	AC023492	AC023492 Homo sapi
c	30	21	53.8	110000	1	BA000028_09	Continuation (10 o
c	31	21	53.8	110000	1	BA000028_10	Continuation (11 o
	32	21	53.8	110000	14	CT005269_05	Continuation (6 of
c	33	21	53.8	165789	14	CR759741	CR759741 Danio rer
c	34	21	53.8	198005	9	AC144762	AC144762 Mus muscu
c	35	21	53.8	235411	14	AC084066	AC084066 Mus muscu
	36	20.8	53.3	110000	15	CR382131_25	Continuation (26 o
	37	20.8	53.3	126065	14	AC144677	AC144677 Rattus no
	38	20.8	53.3	186945	9	AC105173	AC105173 Mus muscu
	39	20.8	53.3	230420	9	AC151413	AC151413 Mus muscu
	40	20.8	53.3	231474	9	AC150898	AC150898 Mus muscu
	41	20.8	53.3	246759	14	AC145759	AC145759 Microcebu
	42	20.8	53.3	254107	14	AC158704	AC158704 Microcebu
c	43	20.6	52.8	2762	5	BC090819	BC090819 Danio rer
	44	20.6	52.8	6410	9	AF319526	AF319526 Mus muscu
	45	20.6	52.8	42014	14	AC018327	AC018327 Drosophil

ALIGNMENTS

RESULT 1
AF364060
LOCUS AF364060 5145 bp DNA linear BCT 03-MAY-2001
DEFINITION Cyanothece sp. PCC 8801 Ca-binding protein gene, complete cds.
ACCESSION AF364060
VERSION AF364060.1 GI:13937121
KEYWORDS .
SOURCE Cyanothece sp. PCC 8801 (Synechococcus sp. RF-1)
ORGANISM Cyanothece sp. PCC 8801
Bacteria; Cyanobacteria; Chroococcales; Cyanothece.
REFERENCE 1 (bases 1 to 5145)
AUTHORS Lin,R.-F., Lee,J.-Y. and Huang,T.-C.
TITLE Ca-binding protein
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 5145)
AUTHORS Lin,R.-F., Lee,J.-Y. and Huang,T.-C.
TITLE Direct Submission
JOURNAL Submitted (25-MAR-2001) Academia Sinica, Institute of Botany, Nankang, Taipei, Taiwan 115, Republic of China
FEATURES Location/Qualifiers
source 1. .5145
/organism="Cyanothece sp. PCC 8801"
/mol_type="genomic DNA"
/strain="RF-1; PCC 8801"
/db_xref="taxon:41431"
/note="synonym: Synechococcus sp. RF-1"
CDS 1830. .4079
/codon_start=1
/transl_table=11
/product="Ca-binding protein"
/protein_id="AAK50055.1"
/db_xref="GI:13937122"
/translation="MAFPSLTFLQNEGFTVSDGYTSQDSTPRILADINADGLDDIVGF
SEDAVYVSYQLADGLFTQYRELTFGQVSNFTGLQGYTDNTYVRTFGDVDGDGDDDLLA
FTPTGIQLSINAGTDLTTDYRNDSAVGFQTAAGVTITAAVDGTNFTDPAQWTNQDETP
RYFTDIDGDGDADLIGFGLDNTYYALGNGDGTFASGEASGFWDGAADQPFFSVDQGWT
SDNSTPRLLGDLNGDGWADIIGFAQEGIDFTAFGGTADEGGVFVALSELDGTFGTPTQ
LTIGGTTDPFEDVWSSQDETPFYLVDIDNSGDLDIVGFGPSSIFYALGDGAGGFAAFV
ELESDAFTAASGFGSNNETPRFFTDLNGDTIADILGFGQSTVSSSENVFATDPTIGFD
ESTTTTIGAPSAFTINDGGWTSQTETTRFLTDVDGDGDDDILGFGLNNVYISVNEGGS
FVEPVSVTITGATGFTSSNGGWSDQNLATRLIGDVNNDGRADIIGFGGGGTVFRALGT

GTGFGTSSTISFTGDTTFTTTNGWTSQDETPRFLVDLNNDGFDDLIGFGGTSVFVSMN
NAGIFGAVEQIEIGGATGFTINNGGWASQTQTPRLLGDITGDGFADIVGFGGASVFVA
EGVGDGTFGDATAITFDGGTQFTIAQGWADNNVTPRELTDIDGDGALDIVGFGGNFVF
ASLNDGSGNFGAVQQLDFDGLSQFNINNGGWANQDTTPRFFGEIDGSAGSELLAFGGN
YIFGADVTA"

ORIGIN

Query Match 61.5%; Score 24; DB 1; Length 5145;
Best Local Similarity 84.4%; Pred. No. 9.1;
Matches 27; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 ATAGCGAGCGTCTAATTCTAGAACACTCACAC 32
||||||| || |||||||| ||| ||||| ||
Db 4430 ATAGCGATCGCCTAATTCTGGAAAACTCAAAC 4461

RESULT 2
AC127144
LOCUS AC127144 281517 bp DNA linear HTG 03-OCT-2002
DEFINITION Rattus norvegicus clone CH230-167K14, *** SEQUENCING IN PROGRESS ***, 4 unordered pieces.
ACCESSION AC127144
VERSION AC127144.2 GI:23096486
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 281517)
AUTHORS Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D., Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 281517)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (14-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 3 (bases 1 to 281517)
AUTHORS Rat Genome Sequencing Consortium.
TITLE Direct Submission
JOURNAL Submitted (03-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT On Sep 18, 2002 this sequence version replaced gi:21745674.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequening reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). As a result, the sequence may extend beyond the ends of the clone and there may be contigs that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.
---------- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
---------- Project Information
Center project name: GGMJ
Center clone name: CH230-167K14
---------- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 218698 bases at least Q40
Consensus quality: 221283 bases at least Q30
Consensus quality: 222924 bases at least Q20
Estimated insert size: 244447; sum-of-contigs estimation
Quality coverage: 3x in Q20 bases; sum-of-contigs estimation
----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 268848: contig of 268848 bp in length
* 268849 268948: gap of unknown length
* 268949 278538: contig of 9590 bp in length
* 278539 278638: gap of unknown length
* 278639 279660: contig of 1022 bp in length
* 279661 279760: gap of unknown length
* 279761 281517: contig of 1757 bp in length.
FEATURES Location/Qualifiers
source 1. .281517
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-167K14"
misc_feature 1. .1155
/note="wgs_contig"
gap 268849. .268948
/estimated_length=unknown
gap 278539. .278638
/estimated_length=unknown

```
     gap             279661. .279760
                     /estimated_length=unknown
ORIGIN

  Query Match            58.5%;  Score 22.8;  DB 14;  Length 281517;
  Best Local Similarity  79.4%;  Pred. No. 37;
  Matches   27;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 ATAGCGAGCGTCTAATTCTAGAACACTCACACGC 34
              ||||||||  | ||| ||| | ||||| |||| ||
Db     100950 ATAGCGAGAGGCTACTTCAAAAACACACACAAGC 100983


RESULT 3
BD010597
LOCUS       BD010597                653 bp    DNA     linear   PAT 31-JAN-2002
DEFINITION  Helicobacter polypeptides and corresponding polynucleotide
            molecules.
ACCESSION   BD010597
VERSION     BD010597.1  GI:18638970
KEYWORDS    JP 2001503637-A/35.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 653)
  AUTHORS   Haas,R., Kleanthous,H., Tomb,J.F., Miller,C., Garawi,A.A.,
            Odenbreit,S. and Meyer,T.
  TITLE     Helicobacter polypeptides and corresponding polynucleotide
            molecules
  JOURNAL   Patent: JP 2001503637-A 35 21-MAR-2001;
            MERIEUX ORAVAX, MAX PLANCK GESELLSCHAFT ZUR FORDERUNG DER
            WISSENSCHAFTEN EV BERLIN, HUMAN GENOME SCIENCES INC
COMMENT     OS   Unidentified
            PN   JP 2001503637-A/35
            PD   21-MAR-2001
            PF   14-NOV-1997 JP 1998522949
            PR   14-NOV-1996 US    08/749051,01-APR-1997 US    08/831309 PR
            01-APR-1997 US    08/834705,01-APR-1997 US    08/833457 PR
            24-JUN-1997 US    08/881227,29-JUL-1997 US    08/902615 PI   RAINER
            HAAS,HAROLD KLEANTHOUS,JEAN FRANCOIS TOMB, PI   CHARLES MILLER,
            PI   AMAL AL GARAWI,STEFAN ODENBREIT,THOMAS MEYER
            PC   C07H21/02,C07K1/00,A61K39/02,A01N37/18
            CC   Strandedness: Single;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers
            FT   source          1. .653
            FT   /organism='Unidentified'.
FEATURES              Location/Qualifiers
     source           1. .653
                      /organism="unidentified"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:32644"
ORIGIN

  Query Match            57.4%;  Score 22.4;  DB 6;  Length 653;
  Best Local Similarity  95.8%;  Pred. No. 47;
  Matches   23;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         16 TTCTAGAACACTCACACGCGTCTT 39
              | ||||||||||||||||||||||
Db        547 TGCTAGAACACTCACACGCGTCTT 570


RESULT 4
AK100414
LOCUS       AK100414                1284 bp    mRNA    linear   PLN 24-JUL-2003
DEFINITION  Oryza sativa (japonica cultivar-group) cDNA clone:J023087M06, full
            insert sequence.
ACCESSION   AK100414
VERSION     AK100414.1  GI:32985623
KEYWORDS    FLI_CDNA; CAP trapper.
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1
  AUTHORS   The Rice Full-Length cDNA Consortium, National Institute of
            Agrobiological Sciences Rice Full-Length cDNA Project Team:,
            Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K.,
            Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I.,
            Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C.,
            Ohtsuki,K., Shishiki,T., Foundation of Advancement of International
            Science Genome Sequencing & Analysis Group:, Otomo,Y., Murakami,K.,
            Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y.,
            Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M.,
            Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J.,
            Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J.,
            Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN:,
            Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S.,
            Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M.,
            Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y.,
            Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
            Yoshino,M. and Hayashizaki,Y.
  TITLE     Collection, mapping, and annotation of over 28,000 cDNA clones from
            japonica rice
  JOURNAL   Science 301 (5631), 376-379 (2003)
  PUBMED    12869764
REFERENCE   2  (bases 1 to 1284)
  AUTHORS   Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K.,
            Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W.,
            Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T.,
            Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K.,
            Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I.,
            Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M.,
            Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M.,
            Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
            Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M.,
            Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A.,
            Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M.,
            Namiki,T., Narikawa,R., Niikura,J., Nishi,K., Nomura,K.,
            Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H.,
            Osato,N., Ota,Y., Otomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K.,
            Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K.,
            Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S.,
            Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y.,
            Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A.,
            Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W.,
            Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and
            Yoshimura,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of
            Agrobiological Sciences, Department of Molecular Genetics, Head of
            Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki
            305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp,
            Tel:81-29-838-7007, Fax:81-29-838-7007)
COMMENT     This clone is one of the 28K full-length cDNA clones from japonica
            rice.
            URL : http://cdna01.dna.affrc.go.jp/cDNA/
            NIAS Rice Full-Length cDNA Project Team: Kikuchi,S., Satoh,K.,
            Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J.,
            Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T.,
            Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and
            Yamamoto,M.
            FAIS Genome Sequencing & Analysis Group: Otomo,Y., Iida,Y.,
            Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M.,
            Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J.,
            Mizuno,K., Narikawa,R., Niikura,J., Oka,M., Ryu,R., Sugano,S.,
            Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S.,
            Yoshimura,A., Matsubara,K. and Murakami,K.
            Genome Exploration Research Group in Riken Genomic Sciences Center
            and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K.,
            Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T.,
            Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K.,
            Hiraoka,T., Hori,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y.,
            Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J.,
```

Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N., Ota,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K., Yasunishi,A. and Hayashizaki,Y.

```
FEATURES             Location/Qualifiers
     source          1. .1284
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="mRNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /clone="J023087M06"
ORIGIN

  Query Match             57.4%;  Score 22.4;  DB 15;  Length 1284;
  Best Local Similarity   81.2%;  Pred. No. 48;
  Matches   26;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          4 GCGAGCGTCTAATTCTAGAACACTCACACGCG 35
              || | |  |||| |||||||||| ||||||||
Db        329 GCTATCTGCTAAATCTAGAACACACACACGCG 360
```

```
RESULT 5
AK065982
LOCUS       AK065982               3491 bp    mRNA    linear   PLN 24-JUL-2003
DEFINITION  Oryza sativa (japonica cultivar-group) cDNA clone:J013045J05, full
            insert sequence.
ACCESSION   AK065982
VERSION     AK065982.1  GI:32976000
KEYWORDS    FLI_CDNA; CAP trapper.
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1
  AUTHORS   The Rice Full-Length cDNA Consortium, National Institute of
            Agrobiological Sciences Rice Full-Length cDNA Project Team:,
            Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K.,
            Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I.,
            Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C.,
            Ohtsuki,K., Shishiki,T., Foundation of Advancement of International
            Science Genome Sequencing & Analysis Group:, Otomo,Y., Murakami,K.,
            Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y.,
            Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M.,
            Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J.,
            Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J.,
            Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN:,
            Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S.,
            Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M.,
            Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y.,
            Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
            Yoshino,M. and Hayashizaki,Y.
  TITLE     Collection, mapping, and annotation of over 28,000 cDNA clones from
            japonica rice
  JOURNAL   Science 301 (5631), 376-379 (2003)
   PUBMED   12869764
REFERENCE   2  (bases 1 to 3491)
  AUTHORS   Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K.,
            Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W.,
            Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T.,
            Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K.,
            Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I.,
            Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M.,
            Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M.,
            Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
            Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M.,
            Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A.,
            Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M.,
            Namiki,T., Narikawa,R., Niikura,J., Nishi,K., Nomura,K.,
            Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H.,
            Osato,N., Ota,Y., Otomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K.,
            Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K.,
            Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S.,
            Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y.,
            Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A.,
            Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W.,
            Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and
            Yoshimura,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of
            Agrobiological Sciences, Department of Molecular Genetics, Head of
            Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki
            305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp,
            Tel:81-29-838-7007, Fax:81-29-838-7007)
COMMENT     This clone is one of the 28K full-length cDNA clones from japonica
            rice.
            URL : http://cdna01.dna.affrc.go.jp/cDNA/
            NIAS Rice Full-Length cDNA Project Team: Kikuchi,S., Satoh,K.,
            Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J.,
            Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T.,
            Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and
            Yamamoto,M.
            FAIS Genome Sequencing & Analysis Group: Otomo,Y., Iida,Y.,
            Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M.,
            Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J.,
            Mizuno,K., Narikawa,R., Niikura,J., Oka,M., Ryu,R., Sugano,S.,
            Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S.,
            Yoshimura,A., Matsubara,K. and Murakami,K.
            Genome Exploration Research Group in Riken Genomic Sciences Center
            and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K.,
            Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T.,
            Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K.,
            Hiraoka,T., Hori,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y.,
            Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J.,
            Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
            Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
            Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N.,
            Ota,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H.,
            Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F.,
            Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K.,
            Yasunishi,A. and Hayashizaki,Y.
FEATURES             Location/Qualifiers
     source          1. .3491
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="mRNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /clone="J013045J05"
ORIGIN

  Query Match             57.4%;  Score 22.4;  DB 15;  Length 3491;
  Best Local Similarity   81.2%;  Pred. No. 50;
  Matches   26;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          4 GCGAGCGTCTAATTCTAGAACACTCACACGCG 35
              || | |  |||| |||||||||| ||||||||
Db        837 GCTATCTGCTAAATCTAGAACACACACACGCG 868
```

```
RESULT 6
AE000531/c
LOCUS       AE000531               10530 bp    DNA     linear   BCT 06-APR-1999
DEFINITION  Helicobacter pylori 26695 section 9 of 134 of the complete genome.
ACCESSION   AE000531 AE000511
VERSION     AE000531.1  GI:2313173
KEYWORDS    .
SOURCE      Helicobacter pylori 26695
  ORGANISM  Helicobacter pylori 26695
            Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
            Helicobacteraceae; Helicobacter.
```

REFERENCE 1 (bases 1 to 10530)
AUTHORS Tomb,J.-F., White,O., Kerlavage,A.R., Clayton,R.A., Sutton,G.G., Fleischmann,R.D., Ketchum,K.A., Klenk,H.P., Gill,S., Dougherty,B.A., Nelson,K., Quackenbush,J., Zhou,L., Kirkness,E.F., Peterson,S., Loftus,B., Richardson,D., Dodson,R., Khalak,H.G., Glodek,A., McKenney,K., Fitzegerald,L.M., Lee,N., Adams,M.D., Hickey,E.K., Berg,D.E., Gocayne,J.D., Utterback,T.R., Peterson,J.D., Kelley,J.M., Karp,P.D., Smith,H.O., Fraser,C.M. and Venter,J.C.
TITLE The complete genome sequence of the gastric pathogen Helicobacter pylori
JOURNAL Nature 388 (6642), 539-547 (1997)
PUBMED 9252185
REFERENCE 2 (bases 1 to 10530)
AUTHORS Tomb,J.-F., White,O., Kerlavage,A.R., Clayton,R.A., Sutton,G.G., Fleischmann,R.D., Ketchum,K.A., Klenk,H.P., Gill,S., Dougherty,B.A., Nelson,K., Quackenbush,J., Zhou,L., Kirkness,E.F., Peterson,S., Loftus,B., Richardson,D., Dodson,R., Khalak,H.G., Glodek,A., McKenney,K., Fitzegerald,L.M., Lee,N., Adams,M.D., Hickey,E.K., Berg,D.E., Gocayne,J.D., Utterback,T.R., Peterson,J.D., Kelley,J.M., Cotton,M.D., Weidman,J.M., Fujii,C., Bowman,C., Watthey,L., Wallin,E., Hayes,W.S., Borodovsky,M., Karp,P.D., Smith,H.O., Fraser,C.M. and Venter,J.C.
TITLE Direct Submission
JOURNAL Submitted (06-AUG-1997) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA
REFERENCE 3 (bases 1 to 10530)
AUTHORS White,O.
TITLE Direct Submission
JOURNAL Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA
FEATURES Location/Qualifiers
source 1. .10530
/organism="Helicobacter pylori 26695"
/mol_type="genomic DNA"
/strain="26695"
/db_xref="taxon:85962"
gene 97. .185
/gene="tRNA-Ser-3"
tRNA 97. .185
/gene="tRNA-Ser-3"
/product="tRNA-Ser"
/note="codon recognized: AGC"
gene 394. .1227
/gene="HP0091"
CDS 394. .1227
/gene="HP0091"
/note="similar to GB:D13968 SP:P34719 PID:303629 percent identity: 50.74; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="type II restriction enzyme R protein (hsdR)"
/protein_id="AAD07162.1"
/db_xref="GI:2313174"
/translation="MKNHLPFDIFLKSLKTSNRTLDFFTDWQKCLKNKNKISIALNHL NFLLGKDTKELKNCIKSLFKEYPKAFNVLNILIAVRDKKDIVLDANGNFYPLYSYFED GEKVYEFIRQTGLERIFCNRNIKDLNDFVFGIEVGLDSNARKNRSGKVMENHLSGLFT NAQLNFKEQVEIREFEDLCQAFGDDIKKFDFVVCGKDKTYFIEANFYTISGSKLNEVA RSYQDLALKFEAFPNYEFIWITDGIGWLDAKNKLQEAYKSVEIYNLSNVNDFISKAQK W"
gene 1221. .2054
/gene="HP0092"
CDS 1221. .2054
/gene="HP0092"
/note="similar to GB:D13968 SP:P34721 PID:303630 percent identity: 55.30; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="type II restriction enzyme M protein (hsdM)"
/protein_id="AAD07163.1"
/db_xref="GI:2313175"
/translation="MVIAHSNEIARPIFKSQDQLFTLYQGDCNEVLPQFENQFDLIFA DPPYFLSNDGLSIQSGKIVSVNKGDWDKEDGINGIDEFNYQWINNAKKALKDTGSLLI SGTYHNIFSLGCVLQKLDFKILNLITWQKTNPPPNFSCRYLTHSAEQIIWARKSRKHK HVFNYEVLKKINNDKQMRDVWSFPAIAPWEKVNGKHPTQKPLALLVRLLLMASDENSL IGDPFSGSSTTGIAANLLKREFIGIEKESEFIKISMDRKIELDARYKEIRSKIKDLNH Q"
gene complement(2074. .2514)
/gene="HP0093"
CDS complement(2074. .2514)
/gene="HP0093"
/note="hypothetical protein; identified by GeneMark; putative"
/codon_start=1
/transl_table=11
/product="H. pylori predicted coding region HP0093"
/protein_id="AAD07169.2"
/db_xref="GI:4580425"
/translation="MILAAKNSVFVHIRRGDYVGIGCQLGIDYQKKALEYMAKRVPNM ELFVFCEDLEFTQNLDLGYPFMDMTTRDKEEEAYWDMLLMQSCQHGIIANSTYSWWAA YLIENPEKIIIGPKHWLFGHENILCKEWVKIESHFEVKSQKYNA"
gene complement(2511. .2978)
/gene="HP0094"
CDS complement(2511. .2978)
/gene="HP0094"
/note="similar to PID:1197639 PID:1197650 percent identity: 27.62; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAD07164.1"
/db_xref="GI:2313176"
/translation="MAFKVVQICGGLGNQMFQYAFAKSLQKHSNTPVLLDITSFDWSD RKMQLELFPIDLPYASAKEIAIAKMQHLPKLVRDALKCMGFDRVSQEIVFEYEPKLLK PSRLTYFFGYFQDPRYFDAISPLIKQTFTLPPPPPKIIRIIIKKRKNISASFL"
gene complement(3077. .3607)
/gene="HP0095"
CDS complement(3077. .3607)
/gene="HP0095"
/note="hypothetical protein; identified by GeneMark; putative"
/codon_start=1
/transl_table=11
/product="H. pylori predicted coding region HP0095"
/protein_id="AAD07170.1"
/db_xref="GI:2313182"
/translation="MPKPKKNTLPCSLSVKMSYFMRFLIKWRTRSLSHKMMTLIQILS ILALASKASEDLEEQLKKIKDYIYRTLNAKIASDVYNRVLILVNEYCTNEELFDKESV KISDLLIQDIQLYALVDEMLKEDKYQVQHTILKGIIKRKYDEAYSLNSEDRILLEYQE RLLEHSHASFSNKKFK"
gene 3680. .4624
/gene="HP0096"
CDS 3680. .4624
/gene="HP0096"
/note="similar to GB:L77117 PID:1591676 percent identity: 31.03; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="phosphoglycerate dehydrogenase"
/protein_id="AAD07165.1"
/db_xref="GI:2313177"
/translation="MKTFKKGVILDAKSVGLKALEVLKEVADFDFYEVTPPSQIVERS IEAEIMVLNKVVITQEVLSQLPKLKLICITATGTDNVDIKSAKALGIEVKNVSAYSTE SVAQHTLACALSLLGRINDYDRYCKSGEYSQSDLFTHISDIKMGLIKGSQWGVIGLGT IGKRVAKLAQAFGAKVVYYSPKDKKEEYERLSLKDLLATSDIISIHAPLNESTRDLIA LKELQSLKDGAILINVGRGGIVNEKDLAEILETKDLYYASDVFVKEPFEKDHAFLNPK IQNKLLLTPHIAWAYSDSLKTLVEKTKENIQDFLASQK"
gene complement(4781. .5494)
/gene="HP0097"
CDS complement(4781. .5494)
/gene="HP0097"
/note="hypothetical protein; identified by GeneMark; putative"

```
                     /codon_start=1
                     /transl_table=11
                     /product="H. pylori predicted coding region HP0097"
                     /protein_id="AAD07171.1"
                     /db_xref="GI:2313183"
                     /translation="MKKIVLVAIALLMSACASYKITPEHVTSYNNGIQVMTSTQAKSK
                     VQLEIAQSKLKGLNESPLVLYVAAQVIEGSPVVFSRKAISVSINQTNLPVLSLRQVMK
                     SSFDFEGILQSFNIAVPTTPIDNVNMITPPMFYYGQGGFLAYNGMMYGGMGMYGPGFG
                     MMMMDDVEEQEVMQESRQALKILAINYLKNNTLNVESKAKGGFVVVDTKNLKTPGVVV
                     VKVFLEDEIHTFKIDISKM"
     gene            complement(5599. .7059)
                     /gene="HP0098"
     CDS             complement(5599. .7059)
                     /gene="HP0098"
                     /note="similar to GB:L42023 SP:P44503 PID:1003078
                     PID:1221997 PID:1204345 percent identity: 32.89;
                     identified by sequence similarity; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="threonine synthase (thrC)"
                     /protein_id="AAD07166.1"
                     /db_xref="GI:2313178"
                     /translation="MPFVPTRSLKEKKIDFIEAILNPNAPKGGLYTLERFETLQWQDC
                     LNLSYNDLVECVFERLGLEIPKNLLASALKRYENFDNPKNPAPIFALNERLFVQELYH
                     GPSLAFKDMALQPLASLFSNLAVGKNEKYLMLVSTSGDTGPATLESLAGMPNVFVVCL
                     YPKDGTSLVQKLQMVTQSASNLKVFGISGDFDDAQNALKNLLKDDDFNEALKACQLKL
                     SVANSVNFGRIAFQIVYHIWGFLELYKKGAINSKEKITLAIPSGNFGNALGAFYAKKM
                     GLNIDKIKVVTNSNDVLREFIETGRYDLTHRSLKQTYSPAMDILKSSNVERALFSLFG
                     FERTLELMQALEEEKFYALKPKELALLQEHFSCASCSDEACLKTIQEVYAEHQYLIDP
                     HTATALNASLKTHEKTLVSATASYEKFPRITLLALNEQKKNDNDKAALETLKNSYNTP
                     DSQRLDDLFERGIKHQEVLKLNEIKSSILLWLESLH"
     gene            7263. .9290
                     /gene="HP0099"
     CDS             7263. .9290
                     /gene="HP0099"
                     /note="similar to GB:L29189 SP:P39216 PID:459689
                     GB:AL009126 percent identity: 32.81; identified by
                     sequence similarity; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="methyl-accepting chemotaxis protein (tlpA)"
                     /protein_id="AAD07167.1"
                     /db_xref="GI:2313179"
                     /translation="MSKGLSIGNKIILCVALIVIVCVSILGVSLNSRVKEILKESALH
                     SMQDSLHFKVNEVQGVLENTYTSMGIVKEMLPKDTKREIKIGLLKNFILANSHVAGVS

  Query Match             57.4%;  Score 22.4;  DB 1;  Length 10530;
  Best Local Similarity   95.8%;  Pred. No. 51;
  Matches   23;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         16 TTCTAGAACACTCACACGCGTCTT 39
              | ||||||||||||||||||||||
Db       3123 TGCTAGAACACTCACACGCGTCTT 3100


RESULT 7
AP008210_304
WPCOMMENT
  Sequence split into 355 fragments  LOCUS AP008210 Accession AP008210
     Fragment Name      Begin      End
     AP008210_000           1   110000
     AP008210_001      100001   210000
     AP008210_002      200001   310000
     AP008210_003      300001   410000
     AP008210_004      400001   510000
     AP008210_005      500001   610000
     AP008210_006      600001   710000
     AP008210_007      700001   810000
     AP008210_008      800001   910000
     AP008210_009      900001  1010000
     AP008210_010     1000001  1110000
     AP008210_011     1100001  1210000
     AP008210_012     1200001  1310000
     AP008210_013     1300001  1410000
     AP008210_014     1400001  1510000
     AP008210_015     1500001  1610000
     AP008210_016     1600001  1710000
     AP008210_017     1700001  1810000
     AP008210_018     1800001  1910000
     AP008210_019     1900001  2010000
     AP008210_020     2000001  2110000
     AP008210_021     2100001  2210000
     AP008210_022     2200001  2310000
     AP008210_023     2300001  2410000
     AP008210_024     2400001  2510000
     AP008210_025     2500001  2610000
     AP008210_026     2600001  2710000
     AP008210_027     2700001  2810000
     AP008210_028     2800001  2910000
     AP008210_029     2900001  3010000
     AP008210_030     3000001  3110000
     AP008210_031     3100001  3210000
     AP008210_032     3200001  3310000
     AP008210_033     3300001  3410000
     AP008210_034     3400001  3510000
     AP008210_035     3500001  3610000
     AP008210_036     3600001  3710000
     AP008210_037     3700001  3810000
     AP008210_038     3800001  3910000
     AP008210_039     3900001  4010000
     AP008210_040     4000001  4110000
     AP008210_041     4100001  4210000
     AP008210_042     4200001  4310000
     AP008210_043     4300001  4410000
     AP008210_044     4400001  4510000
     AP008210_045     4500001  4610000
     AP008210_046     4600001  4710000
     AP008210_047     4700001  4810000
     AP008210_048     4800001  4910000
     AP008210_049     4900001  5010000
     AP008210_050     5000001  5110000
     AP008210_051     5100001  5210000
     AP008210_052     5200001  5310000
     AP008210_053     5300001  5410000
     AP008210_054     5400001  5510000
     AP008210_055     5500001  5610000
     AP008210_056     5600001  5710000
     AP008210_057     5700001  5810000
     AP008210_058     5800001  5910000
     AP008210_059     5900001  6010000
     AP008210_060     6000001  6110000
     AP008210_061     6100001  6210000
     AP008210_062     6200001  6310000
     AP008210_063     6300001  6410000
     AP008210_064     6400001  6510000
     AP008210_065     6500001  6610000
     AP008210_066     6600001  6710000
     AP008210_067     6700001  6810000
     AP008210_068     6800001  6910000
     AP008210_069     6900001  7010000
     AP008210_070     7000001  7110000
     AP008210_071     7100001  7210000
     AP008210_072     7200001  7310000
     AP008210_073     7300001  7410000
     AP008210_074     7400001  7510000
     AP008210_075     7500001  7610000
     AP008210_076     7600001  7710000
     AP008210_077     7700001  7810000
     AP008210_078     7800001  7910000
     AP008210_079     7900001  8010000
     AP008210_080     8000001  8110000
     AP008210_081     8100001  8210000
     AP008210_082     8200001  8310000
     AP008210_083     8300001  8410000
     AP008210_084     8400001  8510000
     AP008210_085     8500001  8610000
```

AP008210_086	8600001	8710000
AP008210_087	8700001	8810000
AP008210_088	8800001	8910000
AP008210_089	8900001	9010000
AP008210_090	9000001	9110000
AP008210_091	9100001	9210000
AP008210_092	9200001	9310000
AP008210_093	9300001	9410000
AP008210_094	9400001	9510000
AP008210_095	9500001	9610000
AP008210_096	9600001	9710000
AP008210_097	9700001	9810000
AP008210_098	9800001	9910000
AP008210_099	9900001	10010000
AP008210_100	10000001	10110000
AP008210_101	10100001	10210000
AP008210_102	10200001	10310000
AP008210_103	10300001	10410000
AP008210_104	10400001	10510000
AP008210_105	10500001	10610000
AP008210_106	10600001	10710000
AP008210_107	10700001	10810000
AP008210_108	10800001	10910000
AP008210_109	10900001	11010000
AP008210_110	11000001	11110000
AP008210_111	11100001	11210000
AP008210_112	11200001	11310000
AP008210_113	11300001	11410000
AP008210_114	11400001	11510000
AP008210_115	11500001	11610000
AP008210_116	11600001	11710000
AP008210_117	11700001	11810000
AP008210_118	11800001	11910000
AP008210_119	11900001	12010000
AP008210_120	12000001	12110000
AP008210_121	12100001	12210000
AP008210_122	12200001	12310000
AP008210_123	12300001	12410000
AP008210_124	12400001	12510000
AP008210_125	12500001	12610000
AP008210_126	12600001	12710000
AP008210_127	12700001	12810000
AP008210_128	12800001	12910000
AP008210_129	12900001	13010000
AP008210_130	13000001	13110000
AP008210_131	13100001	13210000
AP008210_132	13200001	13310000
AP008210_133	13300001	13410000
AP008210_134	13400001	13510000
AP008210_135	13500001	13610000
AP008210_136	13600001	13710000
AP008210_137	13700001	13810000
AP008210_138	13800001	13910000
AP008210_139	13900001	14010000
AP008210_140	14000001	14110000
AP008210_141	14100001	14210000
AP008210_142	14200001	14310000
AP008210_143	14300001	14410000
AP008210_144	14400001	14510000
AP008210_145	14500001	14610000
AP008210_146	14600001	14710000
AP008210_147	14700001	14810000
AP008210_148	14800001	14910000
AP008210_149	14900001	15010000
AP008210_150	15000001	15110000
AP008210_151	15100001	15210000
AP008210_152	15200001	15310000
AP008210_153	15300001	15410000
AP008210_154	15400001	15510000
AP008210_155	15500001	15610000
AP008210_156	15600001	15710000
AP008210_157	15700001	15810000
AP008210_158	15800001	15910000
AP008210_159	15900001	16010000
AP008210_160	16000001	16110000
AP008210_161	16100001	16210000
AP008210_162	16200001	16310000
AP008210_163	16300001	16410000
AP008210_164	16400001	16510000
AP008210_165	16500001	16610000
AP008210_166	16600001	16710000
AP008210_167	16700001	16810000
AP008210_168	16800001	16910000
AP008210_169	16900001	17010000
AP008210_170	17000001	17110000
AP008210_171	17100001	17210000
AP008210_172	17200001	17310000
AP008210_173	17300001	17410000
AP008210_174	17400001	17510000
AP008210_175	17500001	17610000
AP008210_176	17600001	17710000
AP008210_177	17700001	17810000
AP008210_178	17800001	17910000
AP008210_179	17900001	18010000
AP008210_180	18000001	18110000
AP008210_181	18100001	18210000
AP008210_182	18200001	18310000
AP008210_183	18300001	18410000
AP008210_184	18400001	18510000
AP008210_185	18500001	18610000
AP008210_186	18600001	18710000
AP008210_187	18700001	18810000
AP008210_188	18800001	18910000
AP008210_189	18900001	19010000
AP008210_190	19000001	19110000
AP008210_191	19100001	19210000
AP008210_192	19200001	19310000
AP008210_193	19300001	19410000
AP008210_194	19400001	19510000
AP008210_195	19500001	19610000
AP008210_196	19600001	19710000

Query Match 57.4%; Score 22.4; DB 15; Length 110000;
Best Local Similarity 81.2%; Pred. No. 55;
Matches 26; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy          4 GCGAGCGTCTAATTCTAGAACACTCACACGCG 35
              || | |  |||| |||||||||| ||||||||
Db      63034 GCTATCTGCTAAATCTAGAACACACACACGCG 63065
```

RESULT 8
OSJN00077
LOCUS OSJN00077 166804 bp DNA linear PLN 16-APR-2005
DEFINITION Oryza sativa genomic DNA, chromosome 4, BAC clone: OSJNBa0041A02, complete sequence.
ACCESSION AL606638
VERSION AL606638.2 GI:21740914
KEYWORDS HTG.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1
AUTHORS Feng,Q., Zhang,Y., Hao,P., Wang,S., Fu,G., Huang,Y., Li,Y., Zhu,J., Liu,Y., Hu,X., Jia,P., Zhang,Y., Zhao,Q., Ying,K., Yu,S., Tang,Y., Weng,Q., Zhang,L., Lu,Y., Mu,J., Lu,Y., Zhang,L.S., Yu,Z., Fan,D., Liu,X., Lu,T., Li,C., Wu,Y., Sun,T., Lei,H., Li,T., Hu,H., Guan,J., Wu,M., Zhang,R., Zhou,B., Chen,Z., Chen,L., Jin,Z., Wang,R., Yin,H., Cai,Z., Ren,S., Lv,G., Gu,W., Zhu,G., Tu,Y., Jia,J., Zhang,Y., Chen,J., Kang,H., Chen,X., Shao,C., Sun,Y., Hu,Q., Zhang,X., Zhang,W., Wang,L., Ding,C., Sheng,H., Gu,J., Chen,S., Ni,L., Zhu,F., Chen,W., Lan,L., Lai,Y., Cheng,Z., Gu,M., Jiang,J., Li,J., Hong,G., Xue,Y. and Han,B.
TITLE Sequence and analysis of rice chromosome 4

```
  JOURNAL   Nature 420 (6913), 316-320 (2002)
   PUBMED   12447439
REFERENCE   2
  AUTHORS   Han,B., Feng,Q., Huang,Y.C., Li,Y., Zhu,J.J., Zhao,Q., Hu,X.,
            Liu,Y.L., Mu,J., Yu,Z., Chen,L., Fan,D.L., Weng,Q.J., Zhang,L.,
            Lu,Y.Q., Yu,S.L., Liu,X.H., Lu,T.T., Zhang,Y.J., Lu,Y., Li,C.,
            Li,T., Zhang,Y., Hu,H., Jia,P.X., Qian,Y.M., Ying,K., Zhou,B.,
            Chen,Z.H., Hao,P., Zhang,L., Wu,M., Zhang,R.Q., Guan,J.P., Fu,G.,
            Wang,S.Y., Ren,S.X., Lv,G., Lin,W., Gu,W.Q., Zhu,G.F., Tu,Y.F.,
            Jia,J., Yin,H.F., Zhang,Y., Cai,Z., Chen,J., Kang,H., Chen,X.Y.,
            Shao,C.Y., Sun,Y., Hu,Q.P., Zhang,X.L., Zhang,W., Wang,L.J.,
            Ding,C.W., Sheng,H.H., Gu,J.L., Chen,S.T., Ni,L., Zhu,F.H. and
            Hong,G.F.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-SEP-2001) Han Bin, National Center for Gene Research,
            Chinese Academy of sciences, 500# Cao Bao Road, Shanghai 200233,
            CHINA. E-mail enquiries: bhan@ncgr.ac.cn. Clone requests:
            bhan@ncgr.ac.cn
  REMARK    Oryza sativa japonica (nipponbare) genomic DNA, chromosome 4, BAC
            clone: OSJNBa0041A02.
COMMENT     On Jul 12, 2002 this sequence version replaced gi:15594097.
            Web site: http://www.ncgr.ac.cn
            -------------- Summary Statistics
            Assembly program: phrap
            --------------
            This is a complete sequence.
            Genes were identified by a combination of several methods: Gene
            prediction programs including Fgenesh (http://www.softberry.com/),
            genscan (http://CCR-081.mit.edu/GENSCAN.html), GeneMarkHMM
            (http://genemark.biology.gatech.edu/GeneMark/), tRNAscan-SE (Sean
            Eddy, http://genome.wustl.edu/eddy/tRNAscan-SE/), searches of the
            complete sequence against NCBI none redundant protein database (nr)
            (ftp://ncbi.nlm.nih.gov/blast/db) and the EST database at NCGR.
FEATURES             Location/Qualifiers
     source          1. .166804
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="genomic DNA"
                     /cultivar="japonica"
                     /cultivar="Nipponbare"
                     /sub_species="japonica"
                     /db_xref="taxon:39947"
                     /chromosome="4"
                     /clone="OSJNBa0041A02"
                     /clone_lib="CUGI-OSJNBa"
     gene            4983. .7320
                     /gene="OSJNBa0041A02.1"
     CDS             join(4983. .5309,5356. .6245,6345. .7320)
                     /gene="OSJNBa0041A02.1"
                     /codon_start=1
                     /protein_id="CAE01814.1"
                     /db_xref="GI:32482840"
                     /db_xref="InterPro:IPR007228"
                     /db_xref="InterPro:IPR007321"
                     /db_xref="UniProt/TrEMBL:Q7FAM6"
                     /translation="MARGSALLDGSVLPPSRIVSERQAGLPRRFMPESATGREIVMLG
                     EARPAPDYPGRSIFFLPFAMAGLVPPFSSFFMDVLEFYDLQMAHLTPNAVMTLAIFAH
                     LCEMFIGSVSPPSVVGGCYFLPRGPVLNRYIPCALRKKWDDWKSDWFYTPLADEAHLR
                     LPSQPPVQASSWRAPVDLGDGYDAVLDRLAGLRSQGLTGAMVYGDYLRRRIAPLQRRA
                     RAAWEYIGPEDYIRTHQGVRWDWAPEDFKILIQRVLNLNSVEASLIPHGILPLCSDPD
                     RVSILTIMTTVGASEEQAPKGHGGAGGSRRGDQSTPGGGRASGPRVGGSGSNRPTDAR
                     GKRKLGGTPPPSPPRGGGAARASSRRPEGAAPTSQPEGERKKKRLRKMGGTEPSQGNL
                     ISPPRWSFNRPPRSDVPSRPSRHPKSGQSEAEDPAAAEERRRESDRREAADRLREAEE
                     AAREAARARQAEDTAREEAARARQAEEAAREEAARARQAEEAAREEAARARQTEAMAS
                     FEAAHDKAAGASLGPTPSGDAQATASGAARDEAAGGSPGPAPSGDAQDQTGPGDIPEL
                     SASSGGPSRVAFSPRRLFSSSSAAPLSAEPLLQALAAANTAVLDGPSAQVEALRAERA
                     ELEAAWARVEEGRRSVDAMVEVGRKAHRRHISELEARKAALAEIAREVEEEREAALIA
                     TTAMIEAQDSLRLQHGRWEAELKKKLDAAQGVLDAAAAREQRVSRPEI"
     gene            16811. .17746
                     /gene="OSJNBa0041A02.2"
     CDS             16811. .17746
                     /gene="OSJNBa0041A02.2"
                     /codon_start=1
                     /protein_id="CAE01815.2"
                     /db_xref="GI:38344135"
                     /db_xref="InterPro:IPR007228"
                     /db_xref="UniProt/TrEMBL:Q7X8Y9"
                     /translation="MALEDHAGAVETSLADREAAAAIREATLAAHEAACAEEESALRL
                     REDALAERERALEESEAATQRLADSLSLREAAQEEQAHRNLECVRAERAALEQRAAYL
                     EAREKELDARERSGGAATGESDLVARLAAAEHTVADMQCALDSSTGEAEALRLSGEIG
                     PGMLWDAVFRLDRAGREAGLWKSQTISRSTNLEGLAPHLRRMAWAVQQLPEELEKTIK
                     SSSRDLAQGAVELVLASYQARDPGFSPWMALEEFPPGTEDDARARVRDAANHVVQSFE
                     GSAPRLAFAPSSDEEGDASGEDDGGDEASDPGALE"
     gene            complement(18362. .24517)
                     /gene="OSJNBa0041A02.3"
     CDS             complement(join(18362. .22090,22163. .24517))
                     /gene="OSJNBa0041A02.3"
                     /codon_start=1
                     /protein_id="CAE01816.2"
                     /db_xref="GI:38344136"
                     /db_xref="GOA:Q7X7G0"
                     /db_xref="InterPro:IPR000477"
                     /db_xref="InterPro:IPR001584"
                     /db_xref="InterPro:IPR002156"
                     /db_xref="InterPro:IPR005162"
                     /db_xref="UniProt/TrEMBL:Q7X7G0"
                     /translation="MAEQAKAHDLSPSASGDDGEPNPRHRARTPPPPPRQSPKRGEAL
                     ERVEKSAASPVAGGAGERRDGERRLLVYGDGSTPQGVLQAAGALLRHPPVVPDPESPA
                     QRWLDDVANLVMTAQQRLGAGGRSAATRTSGAATTGSVSSRRRARRAAAAARHSATTP
                     SSAPPTREDQHEEPDARLDIERRRSDRRAPRATEGASSSRASPRHGREDQPSVPPVGG
                     VGCRAFVASLRNVRWPPRFRPTITEKYDGSVNPTEFLQVYTTGIEAAGGDDRVMANFF
                     PMALKGQARGWLMNLPPASVHSWEDLCQQFTMNFQGTYPRPGEEADLHAVQRGDDESL
                     RSYIQRFCQVRNTIPCIPAHAVIYAFRGGVRHNRMLEKIASKEPQTTAELFQLADRVA
                     RKEEAWTWNPSGSGVAASAAPGSAAQTGRRDRRRKKRSVHSGDEGHVLAVEGAPRATR
                     KGRPASDKKKEAGTPSRERPASKWCSVHNTSLHDLADCRAVKNLAERTRKWEEDRRQE
                     RREGKSAAVPSGKRRSEARQKAPAVDIDDGDDDLGFQEPGATIATVDGGACAHVSRRS
                     FKAMRRELLAAAPTHEATRRARWSEVALTFDQTDHPPCVARGGQIAMVVSPTICNVKL
                     GRVLIDGGAALNILSPAAFDAIKAPGMVLQPSQPIIGVTPGHTWPLGHIDLPVTFGGS
                     ANFRTERVNFDVADLSLPYNAVLGRPALVKFMAAVHYAYLLMKMPGPGGPISVRGDLK
                     VALACMEQRADHLAAATKPEGGDERLGTSAPTAPRQRIATCDEVPEDALVSPLRANAD
                     VFAWRPADMPGVPREVIEHRLAVRPGARPVRQKVRRQAPERQAFIREEVARLLEAGFI
                     REVIHPEWLANPVVVPKANGKLRMCIDYTDLNKACPKDPYPLPRIDQIVDSTAGCDLL
                     CFLDAYSGYHQIRMAREDEEKTAFITPVGTYCYTSMPFGLKNAGPTFQRTTRISLGSQ
                     IGRNVEAYVDDLVVKTRNQETLLSDLAETFENLRSARIKLNPDKCVFGVPAGKLLGFL
                     VSARGIEANPEKIRAIERMRPPSKLRDVQCVTGCMAALSRFISRLGEKALPLFKLLKR
                     SGPFTWTEEAEHALTQLKAYLSSPPVLVAPEPNEPLLLYLAATPQVVSAALVVERDED
                     NPHFVHPHSVLTWPGSERGGEAPESDGGLRPLTTGVGPLPACQTVLGAPDPQEGPKAT
                     AGRPHLTPSDPEANPALTRPGKEQGEEAPEPNGGQRPLTTGVGPLPACPTMPGAPDPQ
                     DGPEATKGRPPLSSSDPEVIGTEDKCAPKGHLNEERPGDTAPSREDRPHRRVQRPVYP
                     VSEALRDAKTRYPQAQKMLYAILMASRKLRHYFQAHRVTVVTSYPLGQILHNREGTGR
                     VVKWAIELSEFDLHFEPRHAIKSQALADFVAEWTPAPETVSIPEASTNPSQLPHTTHW
                     VMQFDGSLSLQGAGAGVTLTSPTRDVLRYLVRLDFRATNNMAEYEGLLAGLRVAAGLG
                     IRRLLVLGDSQLVVNQVCKEYRCSDPQMDAYVRQVRRMERHFDGIELRHVPRRDNMIA
                     DELSRLASSRAQTPPGAFEESLTRPSARPDPLGETDAPDRPPRPVGVQASGPEGSAPS
                     SLRLIAWISEIQTYLTDKTLPEDREGSERVQRISKRYVLVEGTLYRRAANGILLNCIP
                     QEQGVELLADIHEGECGAHSASRTLVGKAFRQGFYWPTALNDAVDLVRRCRACQFHAK
                     QIHQPAQALQIIPLSWPFAVWGLDILGPFKRAPGGFEYLYVAIDKFTKWPEAYPVVKI
                     DKHSALKFIKGITARFGVPNRIITDNGTQFTSELFGDYCEDMGIKLCFASPAHPRSNG
                     QVERANAEILKGLKTKTFNILKKHGDSWIEELPAVLWANRTTPSRATGETPFFLVYGA
                     EAVLPSELTLRSPRATMYCEADQDQLRRDDLDYLEERRRRAALRAARYQQSLRRYHQR
                     HVRARSLCVDDLVLRRVQTRAGLSKLSPMWEGPYRVIGVPRPGSVRLATGDGTELPNP
                     WNIEHLRRFYP"
     gene            complement(29590. .30947)
                     /gene="OSJNBa0041A02.4"
     CDS             complement(join(29590. .29859,30492. .30564,30657. .30792,
                     30908. .30947))
                     /gene="OSJNBa0041A02.4"
                     /codon_start=1
                     /protein_id="CAE01817.2"
                     /db_xref="GI:38344137"
                     /db_xref="GOA:Q7X806"
                     /db_xref="InterPro:IPR002048"
                     /db_xref="UniProt/TrEMBL:Q7X806"
                     /translation="MGQAWASLQDKLQGRQWKEKQVRKITDKVFDRLTEDTKKREQEA
                     LTFEEVYIAVLCVYNDINKYLPGPHYDPPSKEKLKAMMNEYDINLDGLLDREEFAEFI
                     RKLTAESLCAISFKLLVTLVAAPALALATKRATEGVPGVGRVVRKVPNSIYASVITLG
                     VVMAQKSTEGVE"
```

```
     gene            complement(33540. .34568)
                     /gene="OSJNBa0041A02.5"
     CDS             complement(join(33540. .33812,34141. .34213,34309. .34420,
                     34529. .34568))
                     /gene="OSJNBa0041A02.5"
                     /codon_start=1
                     /protein_id="CAE01818.2"
                     /db_xref="GI:38344138"
                     /db_xref="UniProt/TrEMBL:Q7X7P8"
                     /translation="MAQFFCRMWDRAQGKLYVVRICDKVFTELSDKNTQLLDISSLHV
                     ATLMVYNSINKQLAGPHKDPPCMNVVAKQMEQYKKETPEGITLEEFRRLIMEWVRKDL
                     RLVLANKAAVAIMAAPLLAVTTKSAGRQVPRVGPAVEKVPTPLLFTVFSVGLMFLQDI
                     RAGKQ"
     gene            39183. .41690
                     /gene="OSJNBa0041A02.6"
     CDS             join(39183. .39236,39369. .41690)
                     /gene="OSJNBa0041A02.6"
                     /codon_start=1
                     /protein_id="CAE01819.2"
                     /db_xref="GI:38344139"
                     /db_xref="InterPro:IPR000873"
                     /db_xref="UniProt/TrEMBL:Q7FAM5"
                     /translation="MTKMLDDITRQTLLRQAKMLSPVLERKLSIQSDHRRTRDDAVPV
                     PVPVPVVKPIASFSPSLHASEKSSRLRTQGATRRRDGRHGQDNYRLLARVASNRSART
                     TVPPPAPRHRQSPERRLDRLASRRSSRAVTPRRGTVDGGNPLSDMRCPRRGDDSSMER
                     GGGSSSSSRSLSREPSSAVQERGRGLHRGASPAAAPRVGAEGGSSTRRLGRLDSGLSV
                     NLVPPPLARRHGSQRGGRGAATTTKLSSSTDVAATIRSSIRPSSREFMERSPRRAGEA
                     ENDRKEGADAASVSKGRPSRGELSSMERGSSSGRSLSREPSSAVQERGRGLHRGASPP
                     AAVRVGAEGSSTRRLERLDSGLSASMVSRRGTPRAGRGASTPKLSSSTDAAAMTRSRI
                     RPNRDLTERSLRRAGEADEDESLRRCRGKGKGKEKGDDDAASVSMGRPSRPPRRALNR
                     INSSSTYSSSSCPPEPTSSTSGSTSSWVPPRDNASSWAPPPPRGNPPSWVPLPPPPGG

  Query Match             57.4%;  Score 22.4;  DB 15;  Length 166804;
  Best Local Similarity   81.2%;  Pred. No. 55;
  Matches   26;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          4 GCGAGCGTCTAATTCTAGAACACTCACACGCG 35
              || | |  |||| |||||||||| ||||||||
Db      96411 GCTATCTGCTAAATCTAGAACACACACACGCG 96442


RESULT 9
AC145082/c
LOCUS       AC145082               168482 bp    DNA     linear   ROD 11-FEB-2005
DEFINITION  Mus musculus chromosome 3, clone RP24-333B11, complete sequence.
ACCESSION   AC145082
VERSION     AC145082.10  GI:59276190
KEYWORDS    HTG.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 168482)
  AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
  TITLE     Mus musculus chromosome 3, clone RP24-333B11
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 168482)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
            Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
            Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
            Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K.,
            Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
            Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
            Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
            Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
            Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., MacLean,C.,
            Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
            Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
            Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
            Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
            Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,
            Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
            Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
            Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
            Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-JUN-2003) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 168482)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
            Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
            Bloom,T., Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J.,
            Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B.,
            DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
            Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D.,
            Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
            Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
            Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,
            Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
            MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,
            McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
            Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
            O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
            Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
            Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
            Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
            Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-DEC-2004) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   4  (bases 1 to 168482)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
            Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
            Bloom,T., Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J.,
            Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B.,
            DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
            Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D.,
            Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
            Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
            Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,
            Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
            MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,
            McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
            Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
            O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
            Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
            Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
            Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
            Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-FEB-2005) Broad Institute of MIT and Harvard, 320
            Charles Street, Cambridge, MA 02141, USA
COMMENT     On Feb 11, 2005 this sequence version replaced gi:56122096.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Broad Institute of MIT and Harvard
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@broad.mit.edu
            -------------- Project Information
                Center project name: L18729
                Center clone name: 333_B_11
            --------------.
FEATURES             Location/Qualifiers
     source          1. .168482
```

```
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="3"
                     /map="3"
                     /clone="RP24-333B11"
                     /clone_lib="RPCI-24 Male Mouse BAC"
     repeat_region   complement(1. .109)
                     /rpt_family="B2_Mm1"
     repeat_region   complement(114. .299)
                     /rpt_family="B2_Mm1"
     repeat_region   complement(1853. .1983)
                     /rpt_family="B1_MM"
     repeat_region   2133. .2312
                     /rpt_family="B3A"
     repeat_region   complement(2368. .2508)
                     /rpt_family="PB1D9"
     repeat_region   complement(2437. .2521)
                     /rpt_family="B4"
     repeat_region   2574. .2594
                     /rpt_family="(TTTAA)n"
     repeat_region   complement(2595. .2779)
                     /rpt_family="B2_Mm1"
     repeat_region   complement(2833. .2930)
                     /rpt_family="RSINE1"
     repeat_region   3315. .3507
                     /rpt_family="B3A"
     repeat_region   complement(3539. .3686)
                     /rpt_family="B1_MM"
     repeat_region   complement(3690. .3835)
                     /rpt_family="B1_MM"
     repeat_region   3857. .3975
                     /rpt_family="PB1D7"
     repeat_region   4017. .4051
                     /rpt_family="(TTTTTG)n"
     repeat_region   complement(4066. .4175)
                     /rpt_family="B2_Mm2"
     repeat_region   complement(4218. .4348)
                     /rpt_family="PB1D9"
     repeat_region   4421. .4457
                     /rpt_family="(TC)n"
     repeat_region   4457. .4501
                     /rpt_family="(TG)n"
     repeat_region   complement(4684. .4795)
                     /rpt_family="B1F"
     repeat_region   complement(4916. .5101)
                     /rpt_family="B4A"
     repeat_region   complement(5250. .5327)
                     /rpt_family="ID_B1"
     repeat_region   complement(5329. .5397)
                     /rpt_family="ID4"
     repeat_region   complement(5751. .5874)
                     /rpt_family="PB1D9"
     repeat_region   complement(6047. .6211)
                     /rpt_family="URR1A"
     repeat_region   complement(6282. .6428)
                     /rpt_family="B1_MM"
     repeat_region   complement(6786. .6970)
                     /rpt_family="B2_Mm1"
     repeat_region   complement(6981. .7117)
                     /rpt_family="Lx4"
     repeat_region   complement(7144. .7252)
                     /rpt_family="Lx5"
     repeat_region   complement(7274. .7466)
                     /rpt_family="B2_Mm2"
     repeat_region   complement(7473. .7566)
                     /rpt_family="B1F"
     repeat_region   complement(7694. .7837)
                     /rpt_family="RSINE1"
     repeat_region   complement(7911. .8043)
                     /rpt_family="B1_MM"
     repeat_region   8204. .8367
                     /rpt_family="MER20"
     repeat_region   complement(8384. .8559)
                     /rpt_family="B2_Mm2"
     repeat_region   8748. .8863
                     /rpt_family="B1F"
     repeat_region   complement(8896. .9032)
                     /rpt_family="B1F"
     repeat_region   9076. .9486
                     /rpt_family="Lx2B"
     repeat_region   complement(9542. .9753)
                     /rpt_family="B3"
     repeat_region   9836. .9982
                     /rpt_family="B1_MM"
     repeat_region   complement(10025. .10038)
                     /rpt_family="RSINE1"
     repeat_region   10039. .10071
                     /rpt_family="(TG)n"
     repeat_region   complement(10072. .10231)
                     /rpt_family="RSINE1"
     repeat_region   complement(10337. .10521)
                     /rpt_family="B2_Mm1"
     repeat_region   complement(10619. .10781)
                     /rpt_family="B3"
     repeat_region   10783. .10816
                     /rpt_family="(TTTTG)n"

  Query Match            56.9%;  Score 22.2;  DB 9;  Length 168482;
  Best Local Similarity  77.1%;  Pred. No. 69;
  Matches   27;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          4 GCGAGCGTCTAATTCTAGAACACTCACACGCGTCT 38
              || ||| || ||||| |||| | ||||| || |||
Db      16815 GCTAGCCTCAAATTCAAGAAGATTCACATGCCTCT 16781


RESULT 10
AC096623/c
LOCUS       AC096623               196824 bp    DNA     linear   ROD 27-FEB-2003
DEFINITION  Mus musculus strain 129/SvJ chromosome 3 clone mgs1-30a13, complete
            sequence.
ACCESSION   AC096623
VERSION     AC096623.31  GI:28570327
KEYWORDS    HTG.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 196824)
  AUTHORS   Fu,Y., Zhou,L. and Roe,B.A.
  TITLE     Mus musculus BAC Clone mgs1-30a13
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 196824)
  AUTHORS   Fu,Y., Zhou,L. and Roe,B.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-SEP-2001) Department Of Chemistry And Biochemistry,
            The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
            OK 73019, USA
REFERENCE   3  (bases 1 to 196824)
  AUTHORS   Fu,Y., Zhou,L. and Roe,B.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-MAY-2002) Department Of Chemistry And Biochemistry,
            The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
            OK 73019, USA
REFERENCE   4  (bases 1 to 196824)
  AUTHORS   Fu,Y., Zhou,L. and Roe,B.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-MAY-2002) Department Of Chemistry And Biochemistry,
            The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
            OK 73019, USA
REFERENCE   5  (bases 1 to 196824)
  AUTHORS   Fu,Y., Zhou,L. and Roe,B.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-FEB-2003) Department Of Chemistry And Biochemistry,
```

```
            The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
            OK 73019, USA
REFERENCE   6  (bases 1 to 196824)
  AUTHORS   Fu,Y., Zhou,L. and Roe,B.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-FEB-2003) Department Of Chemistry And Biochemistry,
            The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
            OK 73019, USA
COMMENT     On Feb 27, 2003 this sequence version replaced gi:28316621.
            -------------- Genome Center
            Center: Department Of Chemistry And Biochemistry
            The University Of Oklahoma
            Center code:UOKNOR
            ----------------------.
FEATURES             Location/Qualifiers
     source          1. .196824
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="129/SvJ"
                     /db_xref="taxon:10090"
                     /chromosome="3"
                     /clone="mgs1-30a13"
                     /clone_lib="Genome Systems mouse ES cell BAC libraries I
                     and II"
ORIGIN

  Query Match             56.9%;  Score 22.2;  DB 9;  Length 196824;
  Best Local Similarity   77.1%;  Pred. No. 69;
  Matches   27;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          4 GCGAGCGTCTAATTCTAGAACACTCACACGCGTCT 38
              || ||| || ||||| |||| | ||||| || |||
Db     187650 GCTAGCCTCAAATTCAAGAAGATTCACATGCCTCT 187616


RESULT 11
BV317977
LOCUS       BV317977                 613 bp    DNA     linear   STS 26-JAN-2005
DEFINITION  S236P6501RD12.T0 AlaskanMalamute Canis familiaris STS genomic,
            sequence tagged site.
ACCESSION   BV317977
VERSION     BV317977.1  GI:57516469
KEYWORDS    STS.
SOURCE      Canis familiaris (dog)
  ORGANISM  Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
            Canis.
REFERENCE   1  (bases 1 to 613)
  AUTHORS   Lindblad-Toh,K.
  TITLE     The genome sequence of Canis familiaris
  JOURNAL   Unpublished (2004)
COMMENT
            Contact: Kerstin Lindblad-Toh
            Whitehead Institute for Biomedical Research, Center for Genome
            Research
            320 Charles Street, Cambridge, MA 02141, USA
            Tel: 6172521477
            Fax: 6172580903
            Email: kerstli@genome.wi.mit.edu
            Primer A: No sequence submitted
            Primer B: No sequence submitted
            STS size: 613
            Protocol:
            WGS-discovery (WGS):
            Paired-end low-coverage whole genome shotgun reads were generated
            from 9 breeds
            (German Shepherd, Rottweiler, Bedlington Terrier, Beagle, Labrador
            Retriever, English
            Shepherd, Italian Greyhound, Alaskan Malamute and the Portuguese
            Water Dog -100,000 each)
            and five other canids (Chinese, Alaskan, Indian and Spanish Gray
            Wolf as well as the
            Californian Coyote).
            The WGS reads were placed uniquely on the CanFam1.0 boxer assembly
            and SNP detection was
            carried out by SSAHA-SNP. 863872 reads were annotated as STSs and
            485941 SNPs were
            annotated with alleles from the boxer and the breed or canid from
            which the particular
            read came. The validation rate for these SNPs was estimated at
            approximately 98%.
            WGA-discovery (WGA) of Boxer/Poodle SNPs:
            A second set of SNPs was generated using a similar methodology
            except that the contigs
            from the 1.5x poodle assembly (Kirkness 2003) were used instead of
            WGS reads. Since this
            sequence lacked base quality scores, arbitrary quality scores of
            phred 40 were assigned
            before the poddle seqeunce was placed uniquely on the CanFam1.0
            boxer assembly and SNP
            detection was carried out by SSAHA-SNP. 1637780 SNPs were annotated
            with alleles from the
            boxer and the poodle. The validation rate for these SNPs was
            estimated at approximately TBD%.
            Internal-WGA-discovery (I-WGA):
            A third set of SNPs were discovered by comparing reads in the WGA
            assembly. SNPs were
            defined as mismatch positions that had a base quality of >= 30 on
            both reads in a region
            that aligned without gaps, and with at most one additional mismatch
            in the ten flanking
            bases. For each allele, at least one additional read had to confirm
            it. 731476 SNPs were
            annotated with alleles between the two boxer alleles. The
            validation rate for these SNPs
            was estimated at approximately TBD%.
FEATURES             Location/Qualifiers
     source          1. .613
                     /organism="Canis familiaris"
                     /mol_type="genomic DNA"
                     /strain="AlaskanMalamute"
                     /db_xref="taxon:9615"
                     /map="+ 20 22-561 45114719-45115265"
                     /clone_lib="AlaskanMalamute"
     STS             <1. .>613
ORIGIN

  Query Match             55.9%;  Score 21.8;  DB 10;  Length 613;
  Best Local Similarity   92.0%;  Pred. No. 89;
  Matches   23;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         10 GTCTAATTCTAGAACACTCACACGC 34
              ||| ||||||||||||||||||| |
Db        458 GTCCAATTCTAGAACACTCACACAC 482


RESULT 12
AC153673/c
LOCUS       AC153673                 84669 bp    DNA     linear   HTG 15-DEC-2004
DEFINITION  Bos taurus clone TAMU-417, WORKING DRAFT SEQUENCE, 5 unordered
            pieces.
ACCESSION   AC153673
VERSION     AC153673.1  GI:56606165
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Bos taurus (cow)
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
            Pecora; Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 84669)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
            Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
            Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
            Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
```

```
            Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
            Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
            Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
            Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
            Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
            Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
            Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
            Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
            Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
            Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
            Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
            Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
            Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
            Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
            Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
            Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
            Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
            Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
            Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
            Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
            Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
            Moser,M., Neal,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N.,
            Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
            Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
            Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
            Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
            Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
            Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
            Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
            Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S.,
            Usmani,K., Vasquez,L., Vera,V., Verduzco,D., Villalon,D.,
            Vinson,R., Waldron,L.M., Wang,Q., Wang,S., Ward-Moore,S.,
            Warren,R., Washington,C., Watlington,S., Williams,G.,
            Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,J.,
            Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstock,G. and Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 84669)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-DEC-2004) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help.tmc.edu
            -------------- Project Information
                Center project name: VUBF
                Center clone name: TAMU-417
            -------------- Summary Statistics
                Sequencing vector: M13;
                Chemistry: Dye-primer Bodipy: 81% of reads
                Chemistry: Dye-terminator Big Dye: 19% of reads
                Assembly program: Phrap; version 0.990329
                Consensus quality: 83533 bases at least Q40
                Consensus quality: 84426 bases at least Q30
                Consensus quality: 84796 bases at least Q20
                Estimated insert size: 84269; sum-of-contigs estimation
                Quality coverage: 3x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 5 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     4695: contig of 4695 bp in length
            *     4696     4795: gap of unknown length
            *     4796    25708: contig of 20913 bp in length
            *    25709    25808: gap of unknown length
            *    25809    51426: contig of 25618 bp in length
            *    51427    51526: gap of unknown length
            *    51527    80965: contig of 29439 bp in length
            *    80966    81065: gap of unknown length
            *    81066    84669: contig of 3604 bp in length.
FEATURES            Location/Qualifiers
     source         1. .84669
                    /organism="Bos taurus"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:9913"
                    /clone="TAMU-417"
     gap            4696. .4795
                    /estimated_length=unknown
     gap            25709. .25808
                    /estimated_length=unknown
     gap            51427. .51526
                    /estimated_length=unknown
     gap            80966. .81065
                    /estimated_length=unknown
ORIGIN

  Query Match           55.9%;  Score 21.8;  DB 14;  Length 84669;
  Best Local Similarity   78.8%;  Pred. No. 1e+02;
  Matches   26;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          6 GAGCGTCTAATTCTAGAACACTCACACGCGTCT 38
              |||| |||| || |||||||| ||   ||||||
Db      31202 GAGCATCTAGTTGTAGAACACCCAGCAGCGTCT 31170


RESULT 13
AP001852
LOCUS       AP001852               168735 bp    DNA     linear   HTG 30-MAY-2000
DEFINITION  Homo sapiens chromosome 18 clone RP11-849I19 map 18q23, WORKING
            DRAFT SEQUENCE, 42 unordered pieces.
ACCESSION   AP001852
VERSION     AP001852.2  GI:8117515
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 168735)
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
            Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Homo sapiens 168,735 genomic DNA of 18q23
  JOURNAL   Published Only in DataBase (2000)
REFERENCE   2  (bases 1 to 168735)
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
            Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-APR-2000) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,
            Japan (E-mail:hattori@gsc.riken.go.jp,
            URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923,
            Fax:81-42-778-9924)
COMMENT     On May 30, 2000 this sequence version replaced gi:7630224.
            -------------- Genome Center
                Center: RIKEN Genomic Sciences Center(GSC)
                Center code: RIKEN
                Web site: http://hgp.gsc.riken.go.jp/
                Contact: hattori@gsc.riken.go.jp
            -------------- Project Information
                Center project name: HumDraft18
                Center clone name: RP11-849I19
            -------------- Summary Statistics
                Sequencing vector: PCR products; 100% of reads
                Chemistry: Dye-terminator ET-amersham; 100% of reads
```

```
    Assembly program: Phrap; version 0.990329
    Consensus quality: 147064 bases at least Q40
    Consensus quality: 157520 bases at least Q30
    Consensus quality: 162179 bases at least Q20
    Insert size: 164635; sum-of-contigs
    Quality coverage: 4.08x in Q20 bases; sum-of-contigs
--------------.
NOTE: This is a 'working draft' sequence. It currently consists of
42 contigs. The true order of the pieces is not known and their
order in this sequence record is arbitrary. Gaps between the
contigs are represented as runs N, but the exact sizes of the gaps
are unknown. This record will be updated with the finished sequence
as soon as it is available and the accession number will be
preserved
         1     12685 contig of  12685 bp in length
     12786     21378 contig of   8593 bp in length
     21479     31812 contig of  10334 bp in length
     31913     40293 contig of   8381 bp in length
     40394     50693 contig of  10300 bp in length
     50794     57517 contig of   6724 bp in length
     57618     65718 contig of   8101 bp in length
     65819     72757 contig of   6939 bp in length
     72858     80449 contig of   7592 bp in length
     80550     87968 contig of   7419 bp in length
     88069     93179 contig of   5111 bp in length
     93280     99470 contig of   6191 bp in length
     99571    105119 contig of   5549 bp in length
    105220    109327 contig of   4108 bp in length
    109428    114708 contig of   5281 bp in length
    114809    118902 contig of   4094 bp in length
    119003    123406 contig of   4404 bp in length
    123507    126415 contig of   2909 bp in length
    126516    129075 contig of   2560 bp in length
    129176    130072 contig of    897 bp in length
    130173    132075 contig of   1903 bp in length
    132176    134810 contig of   2635 bp in length
    134911    136733 contig of   1823 bp in length
    136834    138736 contig of   1903 bp in length
    138837    141319 contig of   2483 bp in length
    141420    143860 contig of   2441 bp in length
    143961    145965 contig of   2005 bp in length
    146066    147716 contig of   1651 bp in length
    147817    149568 contig of   1752 bp in length
    149669    151034 contig of   1366 bp in length
    151135    153329 contig of   2195 bp in length
    153430    155008 contig of   1579 bp in length
    155109    156283 contig of   1175 bp in length
    156384    157915 contig of   1532 bp in length
    158016    159389 contig of   1374 bp in length
    159490    161187 contig of   1698 bp in length
    161288    162633 contig of   1346 bp in length
    162734    163954 contig of   1221 bp in length
    164055    165397 contig of   1343 bp in length
    165498    166501 contig of   1004 bp in length
    166602    167634 contig of   1033 bp in length
    167735    168735 contig of   1001 bp in length
Sequence updated (26-May-2000).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 42 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*        1    12685: contig of 12685 bp in length
*    12686    12785: gap of 100 bp
*    12786    21378: contig of 8593 bp in length
*    21379    21478: gap of 100 bp
*    21479    31812: contig of 10334 bp in length
*    31813    31912: gap of 100 bp
*    31913    40293: contig of 8381 bp in length
*    40294    40393: gap of 100 bp
*    40394    50693: contig of 10300 bp in length
*    50694    50793: gap of 100 bp
*    50794    57517: contig of 6724 bp in length
*    57518    57617: gap of 100 bp
*    57618    65718: contig of 8101 bp in length
*    65719    65818: gap of 100 bp
*    65819    72757: contig of 6939 bp in length
*    72758    72857: gap of 100 bp
*    72858    80449: contig of 7592 bp in length
*    80450    80549: gap of 100 bp
*    80550    87968: contig of 7419 bp in length
*    87969    88068: gap of 100 bp
*    88069    93179: contig of 5111 bp in length
*    93180    93279: gap of 100 bp
*    93280    99470: contig of 6191 bp in length
*    99471    99570: gap of 100 bp
*    99571   105119: contig of 5549 bp in length
*   105120   105219: gap of 100 bp
*   105220   109327: contig of 4108 bp in length
*   109328   109427: gap of 100 bp
*   109428   114708: contig of 5281 bp in length
*   114709   114808: gap of 100 bp
*   114809   118902: contig of 4094 bp in length
*   118903   119002: gap of 100 bp
*   119003   123406: contig of 4404 bp in length
*   123407   123506: gap of 100 bp
*   123507   126415: contig of 2909 bp in length
*   126416   126515: gap of 100 bp
*   126516   129075: contig of 2560 bp in length
*   129076   129175: gap of 100 bp
*   129176   130072: contig of 897 bp in length
*   130073   130172: gap of 100 bp
*   130173   132075: contig of 1903 bp in length
*   132076   132175: gap of 100 bp
*   132176   134810: contig of 2635 bp in length
*   134811   134910: gap of 100 bp
*   134911   136733: contig of 1823 bp in length
*   136734   136833: gap of 100 bp
*   136834   138736: contig of 1903 bp in length
*   138737   138836: gap of 100 bp
*   138837   141319: contig of 2483 bp in length
*   141320   141419: gap of 100 bp
*   141420   143860: contig of 2441 bp in length
*   143861   143960: gap of 100 bp
*   143961   145965: contig of 2005 bp in length
*   145966   146065: gap of 100 bp
*   146066   147716: contig of 1651 bp in length
*   147717   147816: gap of 100 bp
*   147817   149568: contig of 1752 bp in length
*   149569   149668: gap of 100 bp
*   149669   151034: contig of 1366 bp in length
*   151035   151134: gap of 100 bp
*   151135   153329: contig of 2195 bp in length
*   153330   153429: gap of 100 bp
*   153430   155008: contig of 1579 bp in length
*   155009   155108: gap of 100 bp
*   155109   156283: contig of 1175 bp in length
*   156284   156383: gap of 100 bp
*   156384   157915: contig of 1532 bp in length
*   157916   158015: gap of 100 bp
*   158016   159389: contig of 1374 bp in length
*   159390   159489: gap of 100 bp
*   159490   161187: contig of 1698 bp in length
*   161188   161287: gap of 100 bp
*   161288   162633: contig of 1346 bp in length
*   162634   162733: gap of 100 bp
*   162734   163954: contig of 1221 bp in length
*   163955   164054: gap of 100 bp
*   164055   165397: contig of 1343 bp in length
*   165398   165497: gap of 100 bp
*   165498   166501: contig of 1004 bp in length
*   166502   166601: gap of 100 bp
*   166602   167634: contig of 1033 bp in length
```

```
            *   167635   167734: gap of 100 bp
            *   167735   168735: contig of 1001 bp in length.
FEATURES             Location/Qualifiers
     source          1. .168735
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="18"
                     /map="18q23"
                     /clone="RP11-849I19"
     misc_feature    1. .12685
                     /note="assembly_fragment"
     misc_feature    12786. .21378
                     /note="assembly_fragment"
     misc_feature    21479. .31812
                     /note="assembly_fragment"

  Query Match             55.9%;  Score 21.8;  DB 14;  Length 168735;
  Best Local Similarity   78.8%;  Pred. No. 1.1e+02;
  Matches   26;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 ATAGCGAGCGTCTAATTCTAGAACACTCACACG 33
              || || | | ||||||||||||||  || ||||
Db     160399 ATCGCCACCATCTAATTCTAGAACCTTCTCACG 160431


RESULT 14
AC021857/c
LOCUS       AC021857               177575 bp    DNA     linear   HTG 27-MAR-2001
DEFINITION  Homo sapiens chromosome RPCI-11 clone RP11-563B11, WORKING DRAFT
            SEQUENCE, 12 unordered pieces.
ACCESSION   AC021857
VERSION     AC021857.4  GI:9966354
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 177575)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 177575)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JAN-2000) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Sep 4, 2000 this sequence version replaced gi:8887041.

            ----------------- Genome Center -----------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            -------------- Project Information --------------
            Center project name: H_NH0563B11
            -------------- Summary Statistics  --------------
            Sequencing vector: M13; 87%
            Sequencing vector: plasmid; 13%
            Chemistry: Dye-primer ET; 87% of reads
            Chemistry: Dye-terminator Big Dye; 13% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 168674 bases at least Q40
            Consensus quality: 171499 bases at least Q30
            Consensus quality: 173203 bases at least Q20
            Insert size: 154000; agarose-fp
            Insert size: 176475; sum-of-contigs
            Quality coverage: 7.97 in Q20 bases; agarose-fp
            Quality coverage: 7.62 in Q20 bases; sum-of-contigs
            ------------------------------------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 12 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1     1395: contig of 1395 bp in length
            *    1396     1495: gap of unknown length
            *    1496     2801: contig of 1306 bp in length
            *    2802     2901: gap of unknown length
            *    2902     4325: contig of 1424 bp in length
            *    4326     4425: gap of unknown length
            *    4426     5689: contig of 1264 bp in length
            *    5690     5789: gap of unknown length
            *    5790     7692: contig of 1903 bp in length
            *    7693     7792: gap of unknown length
            *    7793    10411: contig of 2619 bp in length
            *   10412    10511: gap of unknown length
            *   10512    12364: contig of 1853 bp in length
            *   12365    12464: gap of unknown length
            *   12465    15226: contig of 2762 bp in length
            *   15227    15326: gap of unknown length
            *   15327    27260: contig of 11934 bp in length
            *   27261    27360: gap of unknown length
            *   27361    44966: contig of 17606 bp in length
            *   44967    45066: gap of unknown length
            *   45067    64260: contig of 19194 bp in length
            *   64261    64360: gap of unknown length
            *   64361   177575: contig of 113215 bp in length.
FEATURES             Location/Qualifiers
     source          1. .177575
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="RPCI-11"
                     /clone="RP11-563B11"
     misc_feature    1. .1395
                     /note="assembly_name:Contig18"
     gap             1396. .1495
                     /estimated_length=unknown
     misc_feature    1496. .2801
                     /note="assembly_name:Contig25"
     gap             2802. .2901
                     /estimated_length=unknown
     misc_feature    2902. .4325
                     /note="assembly_name:Contig27"
     gap             4326. .4425
                     /estimated_length=unknown
     misc_feature    4426. .5689
                     /note="assembly_name:Contig28"
     gap             5690. .5789
                     /estimated_length=unknown
     misc_feature    5790. .7692
                     /note="assembly_name:Contig29"
     gap             7693. .7792
                     /estimated_length=unknown
     misc_feature    7793. .10411
                     /note="assembly_name:Contig30"
     gap             10412. .10511
                     /estimated_length=unknown
     misc_feature    10512. .12364
                     /note="assembly_name:Contig31"
     gap             12365. .12464
                     /estimated_length=unknown
     misc_feature    12465. .15226
                     /note="assembly_name:Contig32"
     gap             15227. .15326
                     /estimated_length=unknown
     misc_feature    15327. .27260
                     /note="assembly_name:Contig33"
     gap             27261. .27360
                     /estimated_length=unknown
     misc_feature    27361. .44966
```

```
                     /note="assembly_name:Contig34
                     clone_end:T7
                     vector_side:left"
     gap             44967. .45066
                     /estimated_length=unknown
     misc_feature    45067. .64260
                     /note="assembly_name:Contig35
                     clone_end:SP6
                     vector_side:left"
     gap             64261. .64360
                     /estimated_length=unknown
     misc_feature    64361. .177575
                     /note="assembly_name:Contig36"
ORIGIN

  Query Match             55.9%;  Score 21.8;  DB 14;  Length 177575;
  Best Local Similarity   78.8%;  Pred. No. 1.1e+02;
  Matches   26;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 ATAGCGAGCGTCTAATTCTAGAACACTCACACG 33
              || || | | ||||||||||||||  || ||||
Db     124496 ATCGCCACCATCTAATTCTAGAACCTTCTCACG 124464


RESULT 15
AC099689/c
LOCUS       AC099689               201519 bp    DNA     linear   PRI 03-SEP-2002
DEFINITION  Homo sapiens chromosome 18, clone RP11-849I19, complete sequence.
ACCESSION   AC099689
VERSION     AC099689.4  GI:22655826
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 201519)
  AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 18, clone RP11-849I19
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 201519)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
            Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
            Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
            Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
            Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Liu,G.,
            MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
            McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
            Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
            Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
            Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
            Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
            Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-NOV-2001) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 201519)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
            Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
            Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
            Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
            Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
            Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
            McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
            Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
            O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
            Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
            Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
            Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
            Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
            Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-AUG-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   4  (bases 1 to 201519)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
            Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
            Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
            Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
            Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
            Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
            McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
            Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
            O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
            Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
            Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
            Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
            Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
            Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-SEP-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Sep 3, 2002 this sequence version replaced gi:22532588.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L13214
                Center clone name: 849_I_19
            --------------.
FEATURES             Location/Qualifiers
     source          1. .201519
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="18"
                     /map="18"
                     /clone="RP11-849I19"
                     /clone_lib="RPCI-11 Human Male BAC"
     repeat_region   1. .255
                     /rpt_family="L1"
     repeat_region   complement(248. .379)
                     /rpt_family="L1PA15-16"
     repeat_region   373. .1421
                     /rpt_family="L1"
     repeat_region   1422. .1915
                     /rpt_family="L1PA16"
     repeat_region   1921. .2512
                     /rpt_family="L1MEc"
     repeat_region   2653. .2928
                     /rpt_family="L1MEc"
     repeat_region   complement(2929. .3505)
                     /rpt_family="MLT1E1"
```

```
repeat_region   3506. .3708
                /rpt_family="L1MEc"
repeat_region   3840. .4264
                /rpt_family="MLT2F"
repeat_region   4271. .4502
                /rpt_family="L1ME3"
repeat_region   5703. .6155
                /rpt_family="L1MB7"
repeat_region   8041. .8164
                /rpt_family="MIR"
repeat_region   10799. .10967
                /rpt_family="(ATGGTG)n"
repeat_region   11369. .11393
                /rpt_family="(ATG)n"
repeat_region   complement(11798. .11892)
                /rpt_family="MLT1J2"
repeat_region   complement(13779. .13888)
                /rpt_family="L2"
repeat_region   17923. .18222
                /rpt_family="AluJb"
repeat_region   19455. .19760
                /rpt_family="AluJo"
repeat_region   20132. .20302
                /rpt_family="AluYa5"
repeat_region   20303. .20488
                /rpt_family="AluSg/x"
repeat_region   20498. .20786
                /rpt_family="AluSx"
repeat_region   20810. .20980
                /rpt_family="FRAM"
repeat_region   complement(23080. .23364)
                /rpt_family="AluJb"
repeat_region   23903. .24105
                /rpt_family="(CCCCAG)n"
repeat_region   24152. .24211
                /rpt_family="(CCCCAG)n"
repeat_region   24244. .24433
                /rpt_family="(CCCCAG)n"
repeat_region   24475. .24653
                /rpt_family="(CCCCAG)n"
repeat_region   25987. .26093
                /rpt_family="Charlie4"
repeat_region   26229. .26331
                /rpt_family="Charlie4"
repeat_region   28009. .28083
                /rpt_family="MIR"
repeat_region   complement(28483. .28723)
                /rpt_family="MIR3"
repeat_region   28821. .28901
                /rpt_family="(TAGA)n"
repeat_region   28906. .29107
                /rpt_family="(TGGA)n"
repeat_region   29122. .29323
                /rpt_family="(TGGA)n"
repeat_region   29338. .29539
                /rpt_family="(TGGA)n"
unsure          29407. .29560
                /note="single clone coverage"
unsure          29407. .29423
                /note="< 30 qual single clone coverage"
repeat_region   29554. .30046
                /rpt_family="(TGGA)n"
repeat_region   30016. .30191
                /rpt_family="(TAGA)n"
repeat_region   30194. .30493
                /rpt_family="AluYa5"
repeat_region   30494. .30816
                /rpt_family="(TAGA)n"
repeat_region   complement(31718. .31919)
                /rpt_family="AluJ"
repeat_region   complement(31933. .32005)
                /rpt_family="AluJ"
repeat_region   33370. .33548
                /rpt_family="(CA)n"
repeat_region   33561. .33649
                /rpt_family="(CA)n"
repeat_region   33720. .33827
                /rpt_family="L2"
repeat_region   34885. .35203
                /rpt_family="L2"

  Query Match             55.9%;  Score 21.8;  DB 8;  Length 201519;
  Best Local Similarity   78.8%;  Pred. No. 1.1e+02;
  Matches   26;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 ATAGCGAGCGTCTAATTCTAGAACACTCACACG 33
              || || | | ||||||||||||||  || ||||
Db      38836 ATCGCCACCATCTAATTCTAGAACCTTCTCACG 38804


Search completed: January 30, 2006, 02:34:22
Job time : 898 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: January 29, 2006, 22:23:56 ; Search time 219 Seconds
(without alignments)
1186.863 Million cell updates/sec

Title: US-10-663-033-REVCOMP4-SEQ3
Perfect score: 39
Sequence: 1 atagcgagcgtctaattctagaacactcacacgcgtctt 39

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 0.1

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_21:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	22.4	57.4	653	2	AAV52043	Aav52043 Helicobac
c	2	20.6	52.8	2409	13	ADS47793	Ads47793 Bacterial
	3	20.2	51.8	402	4	AAL19396	Aal19396 Human bre
	4	20.2	51.8	402	4	AAL09009	Aal09009 Human bre
	5	20.2	51.8	451	11	ACN79225	Acn79225 Breast ca
	6	20.2	51.8	523	4	AAL09684	Aal09684 Human bre
	7	20.2	51.8	561	4	AAK92531	Aak92531 Human cDN
	8	20.2	51.8	561	12	ADL28958	Adl28958 3' end of
	9	20.2	51.8	603	4	AAL18356	Aal18356 Human bre
	10	20.2	51.8	686	11	ACN79884	Acn79884 Breast ca
c	11	20.2	51.8	1142	4	AAS33195	Aas33195 DNA encod
c	12	20.2	51.8	2709	4	AAK94205	Aak94205 Human ful
c	13	20.2	51.8	2709	12	ADL30735	Adl30735 Full leng
c	14	20.2	51.8	2759	4	AAK53167	Aak53167 Human pol
c	15	20.2	51.8	2759	4	AAI59801	Aai59801 Human pol
c	16	20.2	51.8	2844	4	AAK52183	Aak52183 Human pol
c	17	20.2	51.8	2844	4	AAI58015	Aai58015 Human pol
c	18	20.2	51.8	2864	12	ADQ86453	Adq86453 Human tum
c	19	20.2	51.8	2871	13	ADR99114	Adr99114 GALNT6 co
c	20	20.2	51.8	2881	6	ABT07725	Abt07725 Breast ca
c	21	20.2	51.8	2881	11	ACN44271	Acn44271 Human mRN
c	22	20.2	51.8	2881	12	ADN04939	Adn04939 Antipsori
c	23	20.2	51.8	2881	14	AEB22801	Aeb22801 Human col
c	24	20.2	51.8	3230	11	ACN89199	Acn89199 Breast ca
c	25	20.2	51.8	12216	4	AAS33427	Aas33427 DNA encod
c	26	20.2	51.8	46404	11	ACN44270	Acn44270 Human gen
c	27	20	51.3	20	3	AAA09497	Aaa09497 Primer GW
c	28	20	51.3	934	2	AAX14231	Aax14231 H. pylori
	29	20	51.3	1119	3	AAA09495	Aaa09495 H. pylori
c	30	20	51.3	110000	14	ADZ59507_2	Continuation (3 of
	31	19.8	50.8	386	3	AAC01025	Aac01025 Human sec
	32	19.8	50.8	569	5	ABV35629	Abv35629 Human pro
	33	19.8	50.8	569	5	ABV44435	Abv44435 Human pro
	34	19.8	50.8	570	5	ABV14550	Abv14550 Human pro
c	35	19.8	50.8	622	10	ADB57315	Adb57315 Toxicity-
c	36	19.8	50.8	622	10	ADB51871	Adb51871 Primary r
c	37	19.8	50.8	622	13	ADV40531	Adv40531 Rat cardi
	38	19.8	50.8	697	2	AAZ33534	Aaz33534 Human pro
	39	19.8	50.8	731	6	AAS62285	Aas62285 cDNA sequ
c	40	19.8	50.8	790	6	AAL41606	Aal41606 Human col
c	41	19.8	50.8	800	3	AAC99839	Aac99839 Human sec
	42	19.8	50.8	954	6	ABQ54895	Abq54895 Human ova
	43	19.8	50.8	1046	13	ADR25781	Adr25781 Breast ca
	44	19.8	50.8	1062	13	ADS09846	Ads09846 Human the
	45	19.8	50.8	1249	5	ABV23395	Abv23395 Human pro

ALIGNMENTS

RESULT 1
AAV52043
ID AAV52043 standard; DNA; 653 BP.
XX
AC AAV52043;
XX
DT 09-NOV-1998 (first entry)
XX
DE Helicobacter polypeptide GHPO 1273 DNA.
XX
KW GHPO 1273; infection; therapy; diagnosis; vaccine; gastritis; ulcer; ss.
XX
OS Helicobacter pylori.
XX
FH Key Location/Qualifiers
FT CDS 63. .593
FT /*tag= a
XX
PN WO9821225-A1.
XX
PD 22-MAY-1998.
XX
PF 14-NOV-1997; 97WO-US021353.
XX
PR 14-NOV-1996; 96US-00749051.
PR 01-APR-1997; 97US-00831309.
PR 01-APR-1997; 97US-00833457.
PR 01-APR-1997; 97US-00834705.
PR 24-JUN-1997; 97US-00881227.
PR 29-JUL-1997; 97US-00902615.
XX
PA (INMR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.
PA (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Haas R, Kleanthous H, Tomb J, Miller C, Al-Garawi A, Odenbreit S;
PI Meyer T;
XX
DR WPI; 1998-297855/26.
DR P-PSDB; AAW71508.
XX
PT Helicobacter polynucleotide and polypeptide sequences - useful to treat

PT or prevent gastrointestinal infection.
XX
PS Disclosure; Page 174-175; 362pp; English.
XX
CC This polynucleotide encodes Helicobacter pylori polypeptide GHPO 1273
CC (see AAW71508). 85 Polynucleotides (see AAV52009-V52093) encoding claimed
CC Helicobacter polypeptides (see AAW71474-W71558) have been identified in
CC the H. pylori genome using a transposon shuttle mutagenesis method in
CC which TnMax9 mini-blaM was used for insertional mutagenesis of an H.
CC pylori gene library established in Escherichia coli. The invention
CC provides: methods for producing these polypeptides in recombinant host
CC systems, and related expression cassettes, vectors and transformed or
CC transfected host cells; live vaccine vectors that contain polynucleotides
CC of the invention and which can be used to prevent or treat Helicobacter
CC infection; therapeutic and/or prophylactic methods involving
CC administration of polynucleotide molecules (either in naked form or
CC delivered with a delivery vehicle), polypeptides or monospecific
CC antibodies; methods for detecting the presence of Helicobacter in samples
CC using e.g. polynucleotide molecules; and methods for purifying
CC polypeptides of the invention
XX
SQ Sequence 653 BP; 234 A; 126 C; 104 G; 189 T; 0 U; 0 Other;

Query Match 57.4%; Score 22.4; DB 2; Length 653;
Best Local Similarity 95.8%; Pred. No. 11;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy         16 TTCTAGAACACTCACACGCGTCTT 39
              |||||||||||||||||||||||
Db        547 TGCTAGAACACTCACACGCGTCTT 570
```

RESULT 2
ADS47793/c
ID ADS47793 standard; cDNA; 2409 BP.
XX
AC ADS47793;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polynucleotide #2536.
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polynucleotide; gene; ss.
XX
OS Bacteria.
XX
PN US2003233675-A1.
XX
PD 18-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PR 21-FEB-2002; 2002US-0360039P.
XX
PA (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX
DR WPI; 2004-061375/06.
XX
PT New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
PS Claim 1; SEQ ID NO 26223; 122pp; English.
XX
CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polynucleotide used in
CC the scope of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 2409 BP; 741 A; 461 C; 555 G; 652 T; 0 U; 0 Other;

Query Match 52.8%; Score 20.6; DB 13; Length 2409;
Best Local Similarity 74.3%; Pred. No. 88;
Matches 26; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```
Qy          4 GCGAGCGTCTAATTCTAGAACACTCACACGCGTCT 38
              | ||||| |||| ||||| ||| | |||||  | |
Db        534 GAGAGCGCCTAAATCTAGGACAATAACACGTTTAT 500
```

RESULT 3
AAL19396
ID AAL19396 standard; cDNA; 402 BP.
XX
AC AAL19396;
XX
DT 07-DEC-2001 (first entry)
XX
DE Human breast cancer expressed polynucleotide 11853.
XX
KW Human; breast cancer; cell marker; cytostatic; ss.
XX
OS Homo sapiens.
XX
PN WO200151628-A2.
XX
PD 19-JUL-2001.
XX
PF 10-JAN-2001; 2001WO-US000798.
XX
PR 14-JAN-2000; 2000US-0176077P.
PR 14-MAR-2000; 2000US-0189167P.
PR 24-MAR-2000; 2000US-0192099P.
PR 29-MAR-2000; 2000US-0193480P.
PR 15-MAY-2000; 2000US-0205230P.
PR 09-JUN-2000; 2000US-0211315P.
PR 25-JUL-2000; 2000US-0220534P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Lillie J, Xu Y, Wang Y, Steinmann K;
XX
DR WPI; 2001-451856/48.
XX

```
PT   New peptide useful as a marker for the diagnosis of breast cancer.
XX
PS   Claim 1; Page 2106; 3695pp; English.
XX
CC   The invention relates to human breast cancer expressed polynucleotides
CC   (AAL07544-AAL26789) and methods of assessing whether a patient is
CC   afflicted with breast cancer by examining the correlation between the
CC   expression of certain markers and the cancerous state of breast cells.
CC   The polynucleotides and encoded polypeptides are potential markers for
CC   detecting, diagnosing, monitoring, characterising treating and
CC   potentially preventing breast cancer. The polynucleotides and encoded
CC   polypeptides are also useful for isolating compounds with cytostatic
CC   activity
XX
SQ   Sequence 402 BP; 88 A; 99 C; 110 G; 105 T; 0 U; 0 Other;

  Query Match             51.8%;  Score 20.2;  DB 4;  Length 402;
  Best Local Similarity   75.8%;  Pred. No. 96;
  Matches   25;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          2 TAGCGAGCGTCTAATTCTAGAACACTCACACGC 34
              | | |||| |||| ||||||||||   | ||||
Db         62 TTGGGAGCCTCTATTTCTAGAACAGGAAAACGC 94


RESULT 4
AAL09009
ID   AAL09009 standard; cDNA; 402 BP.
XX
AC   AAL09009;
XX
DT   07-DEC-2001  (first entry)
XX
DE   Human breast cancer expressed polynucleotide 1466.
XX
KW   Human; breast cancer; cell marker; cytostatic; ss.
XX
OS   Homo sapiens.
XX
PN   WO200151628-A2.
XX
PD   19-JUL-2001.
XX
PF   10-JAN-2001; 2001WO-US000798.
XX
PR   14-JAN-2000; 2000US-0176077P.
PR   14-MAR-2000; 2000US-0189167P.
PR   24-MAR-2000; 2000US-0192099P.
PR   29-MAR-2000; 2000US-0193480P.
PR   15-MAY-2000; 2000US-0205230P.
PR   09-JUN-2000; 2000US-0211315P.
PR   25-JUL-2000; 2000US-0220534P.
XX
PA   (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI   Lillie J,  Xu Y,  Wang Y,  Steinmann K;
XX
DR   WPI; 2001-451856/48.
XX
PT   New peptide useful as a marker for the diagnosis of breast cancer.
XX
PS   Claim 1; Page 307; 3695pp; English.
XX
CC   The invention relates to human breast cancer expressed polynucleotides
CC   (AAL07544-AAL26789) and methods of assessing whether a patient is
CC   afflicted with breast cancer by examining the correlation between the
CC   expression of certain markers and the cancerous state of breast cells.
CC   The polynucleotides and encoded polypeptides are potential markers for
CC   detecting, diagnosing, monitoring, characterising treating and
CC   potentially preventing breast cancer. The polynucleotides and encoded
CC   polypeptides are also useful for isolating compounds with cytostatic
CC   activity
XX
SQ   Sequence 402 BP; 85 A; 100 C; 114 G; 103 T; 0 U; 0 Other;

  Query Match             51.8%;  Score 20.2;  DB 4;  Length 402;
  Best Local Similarity   75.8%;  Pred. No. 96;
  Matches   25;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          2 TAGCGAGCGTCTAATTCTAGAACACTCACACGC 34
              | | |||| |||| ||||||||||   | ||||
Db         77 TTGGGAGCCTCTATTTCTAGAACAGGAAAACGC 109


RESULT 5
ACN79225
ID   ACN79225 standard; DNA; 451 BP.
XX
AC   ACN79225;
XX
DT   02-DEC-2004  (first entry)
XX
DE   Breast cancer related marker, seq id 375.
XX
KW   Cancer; breast; tumour; cytostatic; marker; detection; therapy; ds.
XX
OS   Homo sapiens.
XX
PN   US2003099974-A1.
XX
PD   29-MAY-2003.
XX
PF   18-JUL-2002; 2002US-00198846.
XX
PR   18-JUL-2001; 2001US-0306220P.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Lillie J,  Xu Y,  Wang Y,  Steinmann K;
XX
DR   WPI; 2003-787014/74.
XX
PT   Novel isolated polypeptide associated with breast cancer, useful for
PT   detecting presence of polypeptide in sample, as a marker for breast
PT   cancer.
XX
PS   Disclosure; SEQ ID NO 375; 36pp; English.
XX
CC   The invention relates to an isolated polypeptide (I) associated with
CC   breast cancer which is encoded by a nucleic acid molecule comprising a
CC   nucleotide sequence (S1). Further disclosed is an antibody that binds to
CC   the polypeptide of the invention. The activity of the polypeptide of the
CC   invention may be described as cytostatic. The antibody is useful for
CC   detecting the presence of (I) in a sample. Nucleic acid molecules of the
CC   invention are useful in the detection of breast tumours. (I) is useful as
CC   a marker for breast cancer and in breast cancer therapy. Sequences given
CC   in records ACN78851-ACN92934 represent nucleic acid markers associated
CC   with breast cancer. Note: The sequence listing does not form part of the
CC   specification but may be obtained in electronic format from the USPTO web
CC   site at seqdata.uspto.gov/sequence.html?DocID=20030099974
XX
SQ   Sequence 451 BP; 98 A; 112 C; 127 G; 113 T; 0 U; 1 Other;

  Query Match             51.8%;  Score 20.2;  DB 11;  Length 451;
  Best Local Similarity   75.8%;  Pred. No. 98;
  Matches   25;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          2 TAGCGAGCGTCTAATTCTAGAACACTCACACGC 34
              | | |||| |||| ||||||||||   | ||||
Db        109 TTGGGAGCCTCTATTTCTAGAACAGGAAAACGC 141


RESULT 6
AAL09684
ID   AAL09684 standard; cDNA; 523 BP.
XX
```

```
AC   AAL09684;
XX
DT   07-DEC-2001  (first entry)
XX
DE   Human breast cancer expressed polynucleotide 2141.
XX
KW   Human; breast cancer; cell marker; cytostatic; ss.
XX
OS   Homo sapiens.
XX
PN   WO200151628-A2.
XX
PD   19-JUL-2001.
XX
PF   10-JAN-2001; 2001WO-US000798.
XX
PR   14-JAN-2000; 2000US-0176077P.
PR   14-MAR-2000; 2000US-0189167P.
PR   24-MAR-2000; 2000US-0192099P.
PR   29-MAR-2000; 2000US-0193480P.
PR   15-MAY-2000; 2000US-0205230P.
PR   09-JUN-2000; 2000US-0211315P.
PR   25-JUL-2000; 2000US-0220534P.
XX
PA   (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI   Lillie J,  Xu Y,  Wang Y,  Steinmann K;
XX
DR   WPI; 2001-451856/48.
XX
PT   New peptide useful as a marker for the diagnosis of breast cancer.
XX
PS   Claim 1; Page 413; 3695pp; English.
XX
CC   The invention relates to human breast cancer expressed polynucleotides
CC   (AAL07544-AAL26789) and methods of assessing whether a patient is
CC   afflicted with breast cancer by examining the correlation between the
CC   expression of certain markers and the cancerous state of breast cells.
CC   The polynucleotides and encoded polypeptides are potential markers for
CC   detecting, diagnosing, monitoring, characterising treating and
CC   potentially preventing breast cancer. The polynucleotides and encoded
CC   polypeptides are also useful for isolating compounds with cytostatic
CC   activity
XX
SQ   Sequence 523 BP; 110 A; 120 C; 151 G; 122 T; 0 U; 20 Other;

  Query Match             51.8%;  Score 20.2;  DB 4;  Length 523;
  Best Local Similarity   75.8%;  Pred. No. 1e+02;
  Matches   25;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          2 TAGCGAGCGTCTAATTCTAGAACACTCACACGC 34
              | | |||| |||| ||||||||||   | ||||
Db         98 TTGGGAGCCTCTATTTCTAGAACAGGAAAACGC 130


RESULT 7
AAK92531
ID   AAK92531 standard; cDNA; 561 BP.
XX
AC   AAK92531;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human cDNA 3'-end sequence, SEQ ID NO: 991.
XX
KW   Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX
OS   Homo sapiens.
XX
PN   EP1130094-A2.
XX
PD   05-SEP-2001.
XX
PF   07-JUL-2000; 2000EP-00114089.
XX
PR   08-JUL-1999;   99JP-00194486.
PR   11-JAN-2000; 2000JP-00118774.
PR   02-MAY-2000; 2000JP-00183765.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T,  Nishikawa T,  Isogai T,  Hayashi K,  Ishii S,  Kawai Y;
PI   Wakamatsu A,  Sugiyama T,  Nagai K,  Kojima S,  Otsuki T,  Koga H;
XX
DR   WPI; 2001-524255/58.
XX
PT   830 Primers useful for synthesizing full length cDNA clones and their use
PT   in genetic manipulation.
XX
PS   Claim 3; SEQ ID NO 991; 1380pp + Sequence Listing; English.
XX
CC   The invention relates to primers for synthesising full length cDNA
CC   clones. 830 cDNA molecules encoding a human protein have been isolated
CC   and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
CC   been determined. Primers for synthesising the full length cDNA are useful
CC   for clarifying the function of the protein encoded by the cDNA. The full
CC   length clones were obtained by construction of full length enriched cDNA
CC   libraries that were synthesised by the oligo-capping method. The primers
CC   enable the production of the full length cDNA easily without any special
CC   methods. The present sequence is the nucleotide sequence of the 3'-end of
CC   a cDNA provided in the invention. Note: The sequence data for this patent
CC   did not form part of the printed specification, but was obtained in CD-
CC   ROM format directly from EPO
XX
SQ   Sequence 561 BP; 132 A; 146 C; 133 G; 146 T; 0 U; 4 Other;

  Query Match             51.8%;  Score 20.2;  DB 4;  Length 561;
  Best Local Similarity   75.8%;  Pred. No. 1e+02;
  Matches   25;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          2 TAGCGAGCGTCTAATTCTAGAACACTCACACGC 34
              | | |||| |||| ||||||||||   | ||||
Db        369 TTGGGAGCCTCTATTTCTAGAACAGGAAAACGC 401


RESULT 8
ADL28958
ID   ADL28958 standard; cDNA; 561 BP.
XX
AC   ADL28958;
XX
DT   20-MAY-2004  (first entry)
XX
DE   3' end of a human cDNA molecule SeqID 991.
XX
KW   human; medicine; signal transduction; glycoprotein; transcription;
KW   oligo-capping method; ss.
XX
OS   Homo sapiens.
XX
PN   EP1396543-A2.
XX
PD   10-MAR-2004.
XX
PF   07-JUL-2000; 2003EP-00025638.
XX
PR   08-JUL-1999;   99JP-00194486.
PR   11-JAN-2000; 2000JP-00118774.
PR   02-MAY-2000; 2000JP-00183865.
PR   07-JUL-2000; 2000EP-00114089.
XX
PA   (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI   Ota T,  Nishikawa T,  Isogai T,  Hayashi K,  Ishii S,  Kawai Y;
PI   Wakamatsu A,  Sugiyama T,  Nagai K,  Kojima S,  Otsuki T,  Koga H;
XX
```

DR WPI; 2004-204755/20.
XX
PT New oligonucleotide primers (830 cDNAs) useful for synthesizing full
PT length human cDNAs.
XX
PS Disclosure; SEQ ID NO 991; 1340pp; English.
XX
CC This invention relates to a novel primers useful for synthesising full
CC length cDNA molecules that encode human proteins. Specifically, it refers
CC to secretory or membrane proteins that are potential therapeutic agents/
CC target molecules in the field of medicine, and in particular genes
CC encoding proteins that are associated with signal transduction,
CC glycoproteins and transcription. The present invention describes a method
CC for efficiently cloning a full length human cDNA from both the 5' and 3'
CC ends using the oligo-capping method. This polynucleotide sequence is the
CC 3' end of a full length human cDNA sequence of the invention.
XX
SQ Sequence 561 BP; 132 A; 146 C; 133 G; 146 T; 0 U; 4 Other;

Query Match 51.8%; Score 20.2; DB 12; Length 561;
Best Local Similarity 75.8%; Pred. No. 1e+02;
Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 2 TAGCGAGCGTCTAATTCTAGAACACTCACACGC 34
|| |||| |||| |||||||||| | ||||
Db 369 TTGGGAGCCTCTATTTCTAGAACAGGAAAACGC 401

RESULT 9
AAL18356
ID AAL18356 standard; cDNA; 603 BP.
XX
AC AAL18356;
XX
DT 07-DEC-2001 (first entry)
XX
DE Human breast cancer expressed polynucleotide 10813.
XX
KW Human; breast cancer; cell marker; cytostatic; ss.
XX
OS Homo sapiens.
XX
PN WO200151628-A2.
XX
PD 19-JUL-2001.
XX
PF 10-JAN-2001; 2001WO-US000798.
XX
PR 14-JAN-2000; 2000US-0176077P.
PR 14-MAR-2000; 2000US-0189167P.
PR 24-MAR-2000; 2000US-0192099P.
PR 29-MAR-2000; 2000US-0193480P.
PR 15-MAY-2000; 2000US-0205230P.
PR 09-JUN-2000; 2000US-0211315P.
PR 25-JUL-2000; 2000US-0220534P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Lillie J, Xu Y, Wang Y, Steinmann K;
XX
DR WPI; 2001-451856/48.
XX
PT New peptide useful as a marker for the diagnosis of breast cancer.
XX
PS Claim 1; Page 1925; 3695pp; English.
XX
CC The invention relates to human breast cancer expressed polynucleotides
CC (AAL07544-AAL26789) and methods of assessing whether a patient is
CC afflicted with breast cancer by examining the correlation between the
CC expression of certain markers and the cancerous state of breast cells.
CC The polynucleotides and encoded polypeptides are potential markers for
CC detecting, diagnosing, monitoring, characterising treating and
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cytostatic
CC activity
XX
SQ Sequence 603 BP; 160 A; 145 C; 140 G; 158 T; 0 U; 0 Other;

Query Match 51.8%; Score 20.2; DB 4; Length 603;
Best Local Similarity 75.8%; Pred. No. 1e+02;
Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 2 TAGCGAGCGTCTAATTCTAGAACACTCACACGC 34
|| |||| |||| |||||||||| | ||||
Db 255 TTGGGAGCCTCTATTTCTAGAACAGGAAAACGC 287

RESULT 10
ACN79884
ID ACN79884 standard; DNA; 686 BP.
XX
AC ACN79884;
XX
DT 02-DEC-2004 (first entry)
XX
DE Breast cancer related marker, seq id 1034.
XX
KW Cancer; breast; tumour; cytostatic; marker; detection; therapy; ds.
XX
OS Homo sapiens.
XX
PN US2003099974-A1.
XX
PD 29-MAY-2003.
XX
PF 18-JUL-2002; 2002US-00198846.
XX
PR 18-JUL-2001; 2001US-0306220P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Lillie J, Xu Y, Wang Y, Steinmann K;
XX
DR WPI; 2003-787014/74.
XX
PT Novel isolated polypeptide associated with breast cancer, useful for
PT detecting presence of polypeptide in sample, as a marker for breast
PT cancer.
XX
PS Disclosure; SEQ ID NO 1034; 36pp; English.
XX
CC The invention relates to an isolated polypeptide (I) associated with
CC breast cancer which is encoded by a nucleic acid molecule comprising a
CC nucleotide sequence (S1). Further disclosed is an antibody that binds to
CC the polypeptide of the invention. The activity of the polypeptide of the
CC invention may be described as cytostatic. The antibody is useful for
CC detecting the presence of (I) in a sample. Nucleic acid molecules of the
CC invention are useful in the detection of breast tumours. (I) is useful as
CC a marker for breast cancer and in breast cancer therapy. Sequences given
CC in records ACN78851-ACN92934 represent nucleic acid markers associated
CC with breast cancer. Note: The sequence listing does not form part of the
CC specification but may be obtained in electronic format from the USPTO web
CC site at seqdata.uspto.gov/sequence.html?DocID=20030099974
XX
SQ Sequence 686 BP; 152 A; 147 C; 174 G; 154 T; 0 U; 59 Other;

Query Match 51.8%; Score 20.2; DB 11; Length 686;
Best Local Similarity 75.8%; Pred. No. 1.1e+02;
Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 2 TAGCGAGCGTCTAATTCTAGAACACTCACACGC 34
|| |||| |||| |||||||||| | ||||
Db 220 TTGGGAGCCTCTATTTCTAGAACAGGAAAACGC 252

RESULT 11

```
AAS33195/c
ID   AAS33195 standard; cDNA; 1142 BP.
XX
AC   AAS33195;
XX
DT   04-DEC-2001  (first entry)
XX
DE   DNA encoding human secreted protein, Seq ID No 154.
XX
KW   Immunomodulatory; human immunodeficiency virus; HIV; anaemia; angina;
KW   rheumatoid arthritis; antiarteriosclerotic; cardiant; vascular;
KW   cerebroprotective; thrombolytic; antimicrobial; ophthalmological;
KW   cytostatic; Alzheimer's disease; Parkinson's disease; human; cancer;
KW   multiple sclerosis; cancer; hyperproliferative disorder; infection;
KW   Gaucher's disease; neurological disease; cerebrovascular disorder;
KW   thrombosis; wound healing; ss.
XX
OS   Homo sapiens.
XX
PN   WO200155326-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US001347.
XX
PR   31-JAN-2000; 2000US-0179065P.
PR   04-FEB-2000; 2000US-0180628P.
PR   24-FEB-2000; 2000US-0184664P.
PR   02-MAR-2000; 2000US-0186350P.
PR   16-MAR-2000; 2000US-0189874P.
PR   17-MAR-2000; 2000US-0190076P.
PR   18-APR-2000; 2000US-0198123P.
PR   19-MAY-2000; 2000US-0205515P.
PR   07-JUN-2000; 2000US-0209467P.
PR   28-JUN-2000; 2000US-0214886P.
PR   30-JUN-2000; 2000US-0215135P.
PR   07-JUL-2000; 2000US-0216647P.
PR   07-JUL-2000; 2000US-0216880P.
PR   11-JUL-2000; 2000US-0217487P.
PR   11-JUL-2000; 2000US-0217496P.
PR   14-JUL-2000; 2000US-0218290P.
PR   26-JUL-2000; 2000US-0220963P.
PR   26-JUL-2000; 2000US-0220964P.
PR   14-AUG-2000; 2000US-0224518P.
PR   14-AUG-2000; 2000US-0224519P.
PR   14-AUG-2000; 2000US-0225213P.
PR   14-AUG-2000; 2000US-0225214P.
PR   14-AUG-2000; 2000US-0225266P.
PR   14-AUG-2000; 2000US-0225267P.
PR   14-AUG-2000; 2000US-0225268P.
PR   14-AUG-2000; 2000US-0225270P.
PR   14-AUG-2000; 2000US-0225447P.
PR   14-AUG-2000; 2000US-0225757P.
PR   14-AUG-2000; 2000US-0225758P.
PR   14-AUG-2000; 2000US-0225759P.
PR   18-AUG-2000; 2000US-0226279P.
PR   22-AUG-2000; 2000US-0226681P.
PR   22-AUG-2000; 2000US-0226868P.
PR   22-AUG-2000; 2000US-0227182P.
PR   23-AUG-2000; 2000US-0227009P.
PR   30-AUG-2000; 2000US-0228924P.
PR   01-SEP-2000; 2000US-0229287P.
PR   01-SEP-2000; 2000US-0229343P.
PR   01-SEP-2000; 2000US-0229344P.
PR   01-SEP-2000; 2000US-0229345P.
PR   05-SEP-2000; 2000US-0229509P.
PR   05-SEP-2000; 2000US-0229513P.
PR   06-SEP-2000; 2000US-0230437P.
PR   06-SEP-2000; 2000US-0230438P.
PR   08-SEP-2000; 2000US-0231242P.
PR   08-SEP-2000; 2000US-0231243P.
PR   08-SEP-2000; 2000US-0231244P.
PR   08-SEP-2000; 2000US-0231413P.
PR   08-SEP-2000; 2000US-0231414P.
PR   08-SEP-2000; 2000US-0232080P.
PR   08-SEP-2000; 2000US-0232081P.
PR   12-SEP-2000; 2000US-0231968P.
PR   14-SEP-2000; 2000US-0232397P.
PR   14-SEP-2000; 2000US-0232398P.
PR   14-SEP-2000; 2000US-0232399P.
PR   14-SEP-2000; 2000US-0232400P.
PR   14-SEP-2000; 2000US-0232401P.
PR   14-SEP-2000; 2000US-0233063P.
PR   14-SEP-2000; 2000US-0233064P.
PR   14-SEP-2000; 2000US-0233065P.
PR   21-SEP-2000; 2000US-0234223P.
PR   21-SEP-2000; 2000US-0234274P.
PR   25-SEP-2000; 2000US-0234997P.
PR   25-SEP-2000; 2000US-0234998P.
PR   26-SEP-2000; 2000US-0235484P.
PR   27-SEP-2000; 2000US-0235834P.
PR   27-SEP-2000; 2000US-0235836P.
PR   29-SEP-2000; 2000US-0236327P.
PR   29-SEP-2000; 2000US-0236367P.
PR   29-SEP-2000; 2000US-0236368P.
PR   29-SEP-2000; 2000US-0236369P.
PR   29-SEP-2000; 2000US-0236370P.
PR   02-OCT-2000; 2000US-0236802P.
PR   02-OCT-2000; 2000US-0237037P.
PR   02-OCT-2000; 2000US-0237038P.
PR   02-OCT-2000; 2000US-0237039P.
PR   02-OCT-2000; 2000US-0237040P.
PR   13-OCT-2000; 2000US-0239935P.
PR   13-OCT-2000; 2000US-0239937P.
PR   20-OCT-2000; 2000US-0240960P.
PR   20-OCT-2000; 2000US-0241221P.
PR   20-OCT-2000; 2000US-0241785P.
PR   20-OCT-2000; 2000US-0241786P.
PR   20-OCT-2000; 2000US-0241787P.
PR   20-OCT-2000; 2000US-0241808P.
PR   20-OCT-2000; 2000US-0241809P.
PR   20-OCT-2000; 2000US-0241826P.
PR   01-NOV-2000; 2000US-0244617P.
PR   08-NOV-2000; 2000US-0246474P.
PR   08-NOV-2000; 2000US-0246475P.
PR   08-NOV-2000; 2000US-0246476P.
PR   08-NOV-2000; 2000US-0246477P.
PR   08-NOV-2000; 2000US-0246478P.
PR   08-NOV-2000; 2000US-0246523P.
PR   08-NOV-2000; 2000US-0246524P.
PR   08-NOV-2000; 2000US-0246525P.
PR   08-NOV-2000; 2000US-0246526P.
PR   08-NOV-2000; 2000US-0246527P.
PR   08-NOV-2000; 2000US-0246528P.
PR   08-NOV-2000; 2000US-0246532P.
PR   08-NOV-2000; 2000US-0246609P.
PR   08-NOV-2000; 2000US-0246610P.
PR   08-NOV-2000; 2000US-0246611P.
PR   08-NOV-2000; 2000US-0246613P.
PR   17-NOV-2000; 2000US-0249207P.
PR   17-NOV-2000; 2000US-0249208P.
PR   17-NOV-2000; 2000US-0249209P.
PR   17-NOV-2000; 2000US-0249210P.
PR   17-NOV-2000; 2000US-0249211P.
PR   17-NOV-2000; 2000US-0249212P.
PR   17-NOV-2000; 2000US-0249213P.
PR   17-NOV-2000; 2000US-0249214P.
PR   17-NOV-2000; 2000US-0249215P.
PR   17-NOV-2000; 2000US-0249216P.
PR   17-NOV-2000; 2000US-0249217P.
PR   17-NOV-2000; 2000US-0249218P.
PR   17-NOV-2000; 2000US-0249244P.
PR   17-NOV-2000; 2000US-0249245P.
PR   17-NOV-2000; 2000US-0249264P.
PR   17-NOV-2000; 2000US-0249265P.
PR   17-NOV-2000; 2000US-0249297P.
```

```
PR   17-NOV-2000; 2000US-0249299P.
PR   17-NOV-2000; 2000US-0249300P.
PR   01-DEC-2000; 2000US-0250160P.
PR   01-DEC-2000; 2000US-0250391P.
PR   05-DEC-2000; 2000US-0251030P.
PR   05-DEC-2000; 2000US-0251988P.
PR   05-DEC-2000; 2000US-0256719P.
PR   06-DEC-2000; 2000US-0251479P.
PR   08-DEC-2000; 2000US-0251856P.
PR   08-DEC-2000; 2000US-0251868P.
PR   08-DEC-2000; 2000US-0251869P.
PR   08-DEC-2000; 2000US-0251989P.
PR   08-DEC-2000; 2000US-0251990P.
PR   11-DEC-2000; 2000US-0254097P.
PR   05-JAN-2001; 2001US-0259678P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-451931/48.
DR   P-PSDB; AAU20486.
XX
PT   New nucleic acids and polypeptides, useful for diagnosing, preventing or
PT   treating medical conditions.
XX
PS   Claim 1; SEQ ID NO 154; 753pp; English.
XX
CC   The invention relates to novel isolated nucleic acid molecules (I)
CC   encoding human secreted proteins (II). (I) and (II) are used to prevent,
CC   treat or ameliorate a medical condition in e.g. humans, mice, rabbits,
CC   goats, horses, cats, dogs, chickens or sheep. (I) and (II) may be used in
CC   the prevention, treatment and diagnosis of diseases associated with
CC   inappropriate expression of secreted proteins. (I) and complementary
CC   sequences may also be used as DNA probes in diagnostic assays (e.g.
CC   polymerase chain reactions (PCR)) to detect and quantitate the presence
CC   of similar nucleic acid sequences in samples, and so which patients may
CC   be in need of restorative therapy. (II) may also be used as antigens in
CC   the production of antibodies and in assays to identify modulators
CC   (agonists and antagonists) of the expression and activity of the secreted
CC   proteins. The anti-(II) antibodies and antagonists may also be used to
CC   down regulate expression and activity of (II). The anti-(II) antibodies
CC   may also be used as diagnostic agents for detecting the presence of (II)
CC   in samples (e.g. by enzyme linked immunosorbant assay (ELISA)). The
CC   disorders include for example: immune/autoimmune diseases (e.g. HIV
CC   (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis
CC   and multiple sclerosis), cancers and hyperproliferative disorders (e.g.
CC   melanomas, neoplasms of the breast or liver, Sezary syndrome and
CC   Gaucher's disease), neurological diseases (e.g. Alzheimer's disease,
CC   Parkinson's disease and Charcot-Marie-Tooth disease), cardio-/
CC   cerebrovascular disorders (e.g. cardiac arrest, tachycardia, angina and
CC   thrombosis), infections caused by bacteria, viruses and fungi and ocular
CC   disorders (e.g. corneal infections). (I) and (II), agonists, antagonists
CC   and antibodies can also be used to promote wound healing, maintain organs
CC   before transplantation, and support cell culture of primary tissues.

  Query Match             51.8%;  Score 20.2;  DB 4;  Length 1142;
  Best Local Similarity   75.8%;  Pred. No. 1.2e+02;
  Matches   25;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          2 TAGCGAGCGTCTAATTCTAGAACACTCACACGC 34
              | | |||| |||| ||||||||||   | ||||
Db       1094 TTGGGAGCCTCTATTTCTAGAACAGGAAAACGC 1062


RESULT 12
AAK94205/c
ID   AAK94205 standard; cDNA; 2709 BP.
XX
AC   AAK94205;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human full-length cDNA, SEQ ID NO: 2768.
XX
KW   Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX
OS   Homo sapiens.
XX
PN   EP1130094-A2.
XX
PD   05-SEP-2001.
XX
PF   07-JUL-2000; 2000EP-00114089.
XX
PR   08-JUL-1999;   99JP-00194486.
PR   11-JAN-2000; 2000JP-00118774.
PR   02-MAY-2000; 2000JP-00183765.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T,  Nishikawa T,  Isogai T,  Hayashi K,  Ishii S,  Kawai Y;
PI   Wakamatsu A,  Sugiyama T,  Nagai K,  Kojima S,  Otsuki T,  Koga H;
XX
DR   WPI; 2001-524255/58.
DR   P-PSDB; AAM93285.
XX
PT   830 Primers useful for synthesizing full length cDNA clones and their use
PT   in genetic manipulation.
XX
PS   Claim 8; SEQ ID NO 2768; 1380pp + Sequence Listing; English.
XX
CC   The invention relates to primers for synthesising full length cDNA
CC   clones. 830 cDNA molecules encoding a human protein have been isolated
CC   and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
CC   been determined. Primers for synthesising the full length cDNA are useful
CC   for clarifying the function of the protein encoded by the cDNA. The full
CC   length clones were obtained by construction of full length enriched cDNA
CC   libraries that were synthesised by the oligo-capping method. The primers
CC   enable the production of the full length cDNA easily without any special
CC   methods. The present sequence is a full length human cDNA of the
CC   invention. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in CD-ROM format directly
CC   from EPO
XX
SQ   Sequence 2709 BP; 658 A; 756 C; 760 G; 535 T; 0 U; 0 Other;

  Query Match             51.8%;  Score 20.2;  DB 4;  Length 2709;
  Best Local Similarity   75.8%;  Pred. No. 1.3e+02;
  Matches   25;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          2 TAGCGAGCGTCTAATTCTAGAACACTCACACGC 34
              | | |||| |||| ||||||||||   | ||||
Db       2341 TTGGGAGCCTCTATTTCTAGAACAGGAAAACGC 2309


RESULT 13
ADL30735/c
ID   ADL30735 standard; cDNA; 2709 BP.
XX
AC   ADL30735;
XX
DT   20-MAY-2004  (first entry)
XX
DE   Full length human cDNA clone SeqID 2768.
XX
KW   human; medicine; signal transduction; glycoprotein; transcription;
KW   oligo-capping method; ss; gene.
XX
OS   Homo sapiens.
XX
PN   EP1396543-A2.
XX
PD   10-MAR-2004.
XX
PF   07-JUL-2000; 2003EP-00025638.
```

```
XX
PR   08-JUL-1999;   99JP-00194486.
PR   11-JAN-2000; 2000JP-00118774.
PR   02-MAY-2000; 2000JP-00183865.
PR   07-JUL-2000; 2000EP-00114089.
XX
PA   (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI   Ota T,  Nishikawa T,  Isogai T,  Hayashi K,  Ishii S,  Kawai Y;
PI   Wakamatsu A,  Sugiyama T,  Nagai K,  Kojima S,  Otsuki T,  Koga H;
XX
DR   WPI; 2004-204755/20.
DR   P-PSDB; ADL30736.
XX
PT   New oligonucleotide primers (830 cDNAs) useful for synthesizing full
PT   length human cDNAs.
XX
PS   Example 1; SEQ ID NO 2768; 1340pp; English.
XX
CC   This invention relates to a novel primers useful for synthesising full
CC   length cDNA molecules that encode human proteins. Specifically, it refers
CC   to secretory or membrane proteins that are potential therapeutic agents/
CC   target molecules in the field of medicine, and in particular genes
CC   encoding proteins that are associated with signal transduction,
CC   glycoproteins and transcription. The present invention describes a method
CC   for efficiently cloning a full length human cDNA from both the 5' and 3'
CC   ends using the oligo-capping method. This polynucleotide sequence is a
CC   full length human cDNA clone of the invention.
XX
SQ   Sequence 2709 BP; 658 A; 756 C; 760 G; 535 T; 0 U; 0 Other;

  Query Match             51.8%;  Score 20.2;  DB 12;  Length 2709;
  Best Local Similarity   75.8%;  Pred. No. 1.3e+02;
  Matches   25;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          2 TAGCGAGCGTCTAATTCTAGAACACTCACACGC 34
              | | |||| |||| ||||||||||   | ||||
Db       2341 TTGGGAGCCTCTATTTCTAGAACAGGAAAACGC 2309


RESULT 14
AAK53167/c
ID   AAK53167 standard; cDNA; 2759 BP.
XX
AC   AAK53167;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human polynucleotide SEQ ID NO 2696.
XX
KW   Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW   vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW   tissue growth factor; immunomodulatory; cancer; leukaemia;
KW   nervous system disorder; arthritis; inflammation; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157190-A2.
XX
PD   09-AUG-2001.
XX
PF   05-FEB-2001; 2001WO-US004098.
XX
PR   03-FEB-2000; 2000US-00496914.
PR   27-APR-2000; 2000US-00560875.
PR   20-JUN-2000; 2000US-00598075.
PR   19-JUL-2000; 2000US-00620325.
PR   01-SEP-2000; 2000US-00654936.
PR   15-SEP-2000; 2000US-00663561.
PR   20-OCT-2000; 2000US-00693325.
PR   30-NOV-2000; 2000US-00728422.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Drmanac RT,  Asundi V,  Zhou P,  Xu C,  Cao Y;
PI   Ma Y,  Zhao QA,  Wang D,  Wang J,  Zhang J,  Ren F,  Chen R,  Wang ZW;
PI   Xue AJ,  Yang Y,  Wejhrman T,  Goodrich R;
XX
DR   WPI; 2001-476283/51.
DR   P-PSDB; AAM80034.
XX
PT   Nucleic acids encoding polypeptides with cytokine-like activities, useful
PT   in diagnosis and gene therapy.
XX
PS   Claim 1; Page 4933-4934; 6221pp; English.
XX
CC   The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC   encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
CC   cytokine, cell proliferation or cell differentiation or which may induce
CC   production of other cytokines in other cell populations. The
CC   polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC   peptide therapy. The polypeptides have various cytokine-like activities,
CC   e.g. stem cell growth factor activity, haematopoiesis regulating
CC   activity, tissue growth factor activity, immunomodulatory activity and
CC   activin/inhibin activity and may be useful in the diagnosis and/or
CC   treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC   inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
CC   (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
CC   sequence listing were missing at the time of publication
XX
SQ   Sequence 2759 BP; 662 A; 771 C; 781 G; 545 T; 0 U; 0 Other;

  Query Match             51.8%;  Score 20.2;  DB 4;  Length 2759;
  Best Local Similarity   75.8%;  Pred. No. 1.3e+02;
  Matches   25;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          2 TAGCGAGCGTCTAATTCTAGAACACTCACACGC 34
              | | |||| |||| ||||||||||   | ||||
Db       2400 TTGGGAGCCTCTATTTCTAGAACAGGAAAACGC 2368


RESULT 15
AAI59801/c
ID   AAI59801 standard; cDNA; 2759 BP.
XX
AC   AAI59801;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polynucleotide SEQ ID NO 3790.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia; ss.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US034263.
XX
PR   23-DEC-1999;   99US-00471275.
PR   21-JAN-2000; 2000US-00488725.
PR   25-APR-2000; 2000US-00552317.
PR   20-JUN-2000; 2000US-00598042.
PR   19-JUL-2000; 2000US-00620312.
PR   03-AUG-2000; 2000US-00653450.
PR   14-SEP-2000; 2000US-00662191.
PR   19-OCT-2000; 2000US-00693036.
PR   29-NOV-2000; 2000US-00727344.
XX
```

PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.
DR P-PSDB; AAM40645.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
XX
PS Claim 1; SEQ ID NO 3790; 10078pp; English.
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX
SQ Sequence 2759 BP; 662 A; 771 C; 781 G; 545 T; 0 U; 0 Other;

Query Match 51.8%; Score 20.2; DB 4; Length 2759;
Best Local Similarity 75.8%; Pred. No. 1.3e+02;
Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```
Qy          2 TAGCGAGCGTCTAATTCTAGAACACTCACACGC 34
              | | |||| |||| ||||||||||   | ||||
Db       2400 TTGGGAGCCTCTATTTCTAGAACAGGAAAACGC 2368
```

Search completed: January 30, 2006, 02:04:26
Job time : 222 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: January 30, 2006, 01:24:06 ; Search time 1556 Seconds
(without alignments)
1172.684 Million cell updates/sec

Title: US-10-663-033-REVCOMP4-SEQ3
Perfect score: 39
Sequence: 1 atagcgagcgtctaattctagaacactcacacgcgtctt 39

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 0.1

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_htc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	24.2	62.1	560	10	CNS01IEX	AL145642 Anopheles
	2	22.6	57.9	1264	10	CL473630	CL473630 SAIL_204_
	3	22.6	57.9	1277	10	CL499547	CL499547 SAIL_670_
c	4	22.4	57.4	831	9	AQ871927	AQ871927 nbeb0045B
	5	22.2	56.9	264	11	CNS04LUL	AL296598 Tetraodon
	6	22.2	56.9	850	10	AG136295	AG136295 Pan trogl
c	7	22	56.4	506	1	AI179396	AI179396 EST223098
	8	21.8	55.9	439	9	AQ494887	AQ494887 HS_5195_A
	9	21.8	55.9	635	9	CE120785	CE120785 tigr-gss-
c	10	21.8	55.9	935	7	CO731803	CO731803 JBW027A05
	11	21.8	55.9	1032	10	CL491525	CL491525 SAIL_557_
	12	21.4	54.9	295	1	AW407061	AW407061 UI-HF-BL0
	13	21.4	54.9	325	3	BM851087	BM851087 K-EST0131
	14	21.4	54.9	445	3	BM851189	BM851189 K-EST0131
c	15	21.4	54.9	504	1	AA695264	AA695264 GM02505.5
c	16	21.4	54.9	565	5	BU580658	BU580658 in36e09.x
	17	21.4	54.9	807	10	CW975483	CW975483 AIAA-aaa8
	18	21.4	54.9	863	6	CF660787	CF660787 CcLM09a32
	19	21.2	54.4	479	2	BE162200	BE162200 IL2-HT044
	20	21.2	54.4	598	3	BJ192483	BJ192483 BJ192483
c	21	21.2	54.4	756	8	CX104655	CX104655 BI103L08
	22	21	53.8	437	2	BF808552	BF808552 CM1-CI009
c	23	21	53.8	588	8	CV952286	CV952286 PVrpvb_63
	24	21	53.8	590	10	CE532527	CE532527 tigr-gss-
c	25	21	53.8	895	7	CN017674	CN017674 AGENCOURT
	26	21	53.8	953	10	CL516085	CL516085 SAIL_912_
	27	21	53.8	1138	2	BG852398	BG852398 1024034A1
	28	20.8	53.3	571	10	CNS01ZYZ	AL174932 Tetraodon
c	29	20.8	53.3	668	10	CE629311	CE629311 tigr-gss-
	30	20.8	53.3	699	9	BH965686	BH965686 odh14b02.
c	31	20.8	53.3	744	7	CV673875	CV673875 RET7SJ_23
	32	20.8	53.3	840	9	BH661798	BH661798 BOHWC11TR
c	33	20.8	53.3	841	2	BG367933	BG367933 HVSMEi001
c	34	20.6	52.8	476	9	AQ685782	AQ685782 HS_5547_A
	35	20.6	52.8	504	7	CR537994	CR537994 DKFZp459N
	36	20.6	52.8	665	2	BB410213	BB410213 BB410213
	37	20.6	52.8	775	8	CV991159	CV991159 IpCGFr1_5
c	38	20.6	52.8	841	10	CZ380115	CZ380115 ZMMBF0153
	39	20.6	52.8	993	2	BG334170	BG334170 602461652
	40	20.6	52.8	1011	10	DU015562	DU015562 237797 To
c	41	20.6	52.8	1309	5	BU367266	BU367266 603785993
	42	20.6	52.8	1469	10	AG065308	AG065308 Pan trogl
c	43	20.4	52.3	234	10	CW069414	CW069414 104_319_1
	44	20.4	52.3	295	10	CL262429	CL262429 ZMMBBb061
c	45	20.4	52.3	315	2	BB531263	BB531263 BB531263

ALIGNMENTS

RESULT 1
CNS01IEX
LOCUS CNS01IEX 560 bp DNA linear GSS 12-JUN-2001
DEFINITION Anopheles gambiae GSS SP6 end of clone 10O05 of NotreDame1 library from strain PEST of Anopheles gambiae (African malaria mosquito), genomic survey sequence.
ACCESSION AL145642
VERSION AL145642.1 GI:7003804
KEYWORDS GSS.
SOURCE Anopheles gambiae (African malaria mosquito)
ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae; Anophelinae; Anopheles.
REFERENCE 1 (bases 1 to 560)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
REFERENCE 2 (bases 1 to 560)
AUTHORS Roth,C.W., Brey,P.T., Ke,Z., Collins,F.H. and Weissenbach,J.
TITLE Direct Submission
JOURNAL Submitted (16-FEB-2000) BBMI, Institut Pasteur, 25, rue du Dr. Roux, Paris 75015, France
COMMENT This clone is from an A. gambiae BAC library provided by F.H. Collins and sequenced by Genoscope in collaboration with the Laboratory of Biochem. and Biol. Molec. of Insects, Institut Pasteur.
FEATURES Location/Qualifiers
source 1. .560
/organism="Anopheles gambiae"
/mol_type="genomic DNA"
/strain="PEST"
/db_xref="taxon:7165"
/clone="10O05"
/clone_lib="NotreDame1"
/note="end : SP6"
ORIGIN

Query Match 62.1%; Score 24.2; DB 10; Length 560;
Best Local Similarity 78.4%; Pred. No. 15;
Matches 29; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 3 AGCGAGCGTCTAATTCTAGAACACTCACACGCGTCTT 39

```
              |||  ||||||| || ||||| || ||||||| || |
Db        210 AGCCTGCGTCTACTTATAGAAAACGCACACGCTTCAT 246


RESULT 2
CL473630
LOCUS       CL473630               1264 bp    DNA     linear   GSS 01-APR-2004
DEFINITION  SAIL_204_A08.v2 SAIL Collection Arabidopsis thaliana genomic clone
            SAIL_204_A08.v2, genomic survey sequence.
ACCESSION   CL473630
VERSION     CL473630.1  GI:45939341
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1264)
  AUTHORS   Sessions,A., Burke,E., Presting,G., Aux,G., McElver,J., Patton,D.,
            Dietrich,B., Ho,P., Bacwaden,J., Ko,C., Clarke,J.D., Cotton,D.,
            Bullis,D., Snell,J., Miguel,T., Hutchison,D., Kimmerly,B.,
            Mitzel,T., Katagiri,F., Glazebrook,J., Law,M. and Goff,S.A.
  TITLE     A high-throughput Arabidopsis reverse genetics system
  JOURNAL   Plant Cell 14 (12), 2985-2994 (2002)
   PUBMED   12468722
COMMENT     Contact: Sessions A
            Applied Trait Genetics
            Syngenta Biotechnology Inc.
            3054 Cornwallis Rd., Research Triangle Park, NC 27709, USA
            Email: allen.sessions@syngenta.com
            ABRC Stock Number CS809533; T-DNA left border flanking sequences of
            Syngenta Arabidopsis Insertion Library (SAIL) lines are available
            through the Arabidopsis Biological Resource Center (ABRC).
            Sequences represent a pool of amplified genomic regions and not
            single contiguous sequences.
            Class: TDNA tagged.
FEATURES             Location/Qualifiers
     source          1. .1264
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /ecotype="Columbia"
                     /db_xref="taxon:3702"
                     /clone="SAIL_204_A08.v2"
                     /clone_lib="SAIL Collection"
                     /note="T-DNA left border sequences were isolated using a
                     modified TAIL-PCR strategy"
ORIGIN

  Query Match             57.9%;  Score 22.6;  DB 10;  Length 1264;
  Best Local Similarity   75.7%;  Pred. No. 83;
  Matches   28;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 ATAGCGAGCGTCTAATTCTAGAACACTCACACGCGTC 37
              || ||| |||| |||||| || ||||| | |  ||||
Db         75 ATTGCGGGCGTTTAATTCCAGTACACTTAAAAACGTC 111


RESULT 3
CL499547
LOCUS       CL499547               1277 bp    DNA     linear   GSS 01-APR-2004
DEFINITION  SAIL_670_D02.v1 SAIL Collection Arabidopsis thaliana genomic clone
            SAIL_670_D02.v1, genomic survey sequence.
ACCESSION   CL499547
VERSION     CL499547.1  GI:45996464
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1277)
  AUTHORS   Sessions,A., Burke,E., Presting,G., Aux,G., McElver,J., Patton,D.,
            Dietrich,B., Ho,P., Bacwaden,J., Ko,C., Clarke,J.D., Cotton,D.,
            Bullis,D., Snell,J., Miguel,T., Hutchison,D., Kimmerly,B.,
            Mitzel,T., Katagiri,F., Glazebrook,J., Law,M. and Goff,S.A.
  TITLE     A high-throughput Arabidopsis reverse genetics system
  JOURNAL   Plant Cell 14 (12), 2985-2994 (2002)
   PUBMED   12468722
COMMENT     Contact: Sessions A
            Applied Trait Genetics
            Syngenta Biotechnology Inc.
            3054 Cornwallis Rd., Research Triangle Park, NC 27709, USA
            Email: allen.sessions@syngenta.com
            ABRC Stock Number CS829268; T-DNA left border flanking sequences of
            Syngenta Arabidopsis Insertion Library (SAIL) lines are available
            through the Arabidopsis Biological Resource Center (ABRC).
            Sequences represent a pool of amplified genomic regions and not
            single contiguous sequences.
            Class: TDNA tagged.
FEATURES             Location/Qualifiers
     source          1. .1277
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /ecotype="Columbia"
                     /db_xref="taxon:3702"
                     /clone="SAIL_670_D02.v1"
                     /clone_lib="SAIL Collection"
                     /note="T-DNA left border sequences were isolated using a
                     modified TAIL-PCR strategy"
ORIGIN

  Query Match             57.9%;  Score 22.6;  DB 10;  Length 1277;
  Best Local Similarity   75.7%;  Pred. No. 83;
  Matches   28;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 ATAGCGAGCGTCTAATTCTAGAACACTCACACGCGTC 37
              || ||  |||||||||||| || ||| ||| |  ||||
Db         86 ATTGCACGCGTCTAATTCCAGTACATTCAAAAACGTC 122


RESULT 4
AQ871927/c
LOCUS       AQ871927               831 bp    DNA     linear   GSS 03-NOV-1999
DEFINITION  nbeb0045B04r CUGI Rice BAC Library (EcoRI) Oryza sativa (japonica
            cultivar-group) genomic clone nbeb0045B04r, genomic survey
            sequence.
ACCESSION   AQ871927
VERSION     AQ871927.1  GI:6222378
KEYWORDS    GSS.
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 831)
  AUTHORS   Wing,R.A. and Dean,R.A.
  TITLE     A BAC End Sequencing Framework to Sequence the Rice Genome
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Seq primer: GGAAACAGCTATGACCATG
            Class: BAC ends
            High quality sequence start: 35
            High quality sequence stop: 521.
FEATURES             Location/Qualifiers
     source          1. .831
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="genomic DNA"
                     /cultivar="japonica"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
```

/clone="nbeb0045B04r"
/tissue_type="Leaf"
/lab_host="E. coli DH10B"
/clone_lib="CUGI Rice BAC Library (EcoRI)"
/note="Vector: pBACIndigo; Site_1: EcoRI; Site_2: EcoRI;
Rice is the most important food crop in the world. Half of
the world population, especially those inhabiting highly
populated areas of the humid tropics and subtropics, rely
on rice as their primary source of carbohydrate.
Monocotyledonous rice is a diploid plant (2n=24) with a
haploid genome equivalent of 431 Mbp (Arumuganathan and
Earle, 1991). The relatively small genome of rice, three
times larger than that of Arabidopsis, makes it suitable
for genomic studies. In order to facilitate positional
cloning, physical mapping and genome sequencing of rice,
we have constructed a BAC library from Oryza sativa,
Nipponbare variety using EcoRI as the cloning enzyme. The
library contains 55,296 clones with an average insert size
of 121 Kb providing approximatley 15 haploid genome
equivalents. The deep coverage allows the isolation a
particular sequence with a probability of 99.9 %. Three
high density filters, each containing 18,432 clones
(doubly spotted), represent the whole library for colony
screening and can be requested from the Clemson University
BAC/EST Resource Center (www.genome.clemson.edu)."

ORIGIN

Query Match 57.4%; Score 22.4; DB 9; Length 831;
Best Local Similarity 81.2%; Pred. No. 97;
Matches 26; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy          4 GCGAGCGTCTAATTCTAGAACACTCACACGCG 35
              || | |  |||| |||||||||| ||||||||
Db        716 GCTATCTGCTAAATCTAGAACACACACACGCG 685
```

RESULT 5
CNS04LUL
LOCUS CNS04LUL 264 bp DNA linear GSS 01-SEP-2000
DEFINITION Tetraodon nigroviridis genome survey sequence PUC-Ori end of clone 119A14 of library G from Tetraodon nigroviridis, genomic survey sequence.
ACCESSION AL296598
VERSION AL296598.1 GI:8035178
KEYWORDS GSS; genome survey sequence.
SOURCE Tetraodon nigroviridis
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetradontoidea; Tetraodontidae; Tetraodon.
REFERENCE 1
AUTHORS Roest Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F., Saurin,W. and Weissenbach,J.
TITLE Estimate of human gene number provided by genome-wide analysis using Tetraodon nigroviridis DNA sequence
JOURNAL Nat. Genet. 25 (2), 235-238 (2000)
PUBMED 10835645
REFERENCE 2
AUTHORS Roest Crollius,H., Jaillon,O., Dasilva,C., Ozouf-Costaz,C., Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J.
TITLE Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis
JOURNAL Genome Res. 10 (7), 939-949 (2000)
PUBMED 10899143
REFERENCE 3 (bases 1 to 264)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (12-APR-2000) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/Tetraodon.
FEATURES Location/Qualifiers
source 1. .264
/organism="Tetraodon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"
/clone="119A14"
/clone_lib="G"
/note="Genoscope sequence ID : C0BG119BA07SP1
end : PUC-Ori"
ORIGIN

Query Match 56.9%; Score 22.2; DB 11; Length 264;
Best Local Similarity 88.9%; Pred. No. 1e+02;
Matches 24; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy         13 TAATTCTAGAACACTCACACGCGTCTT 39
                |||||  |||||||| |||||||||| ||
Db         55 TAATTTTAGAACACACACACGCGTATT 81
```

RESULT 6
AG136295
LOCUS AG136295 850 bp DNA linear GSS 04-NOV-2001
DEFINITION Pan troglodytes DNA, clone: PTB-149N08.F, genomic survey sequence.
ACCESSION AG136295
VERSION AG136295.1 GI:16665973
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1
AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE BAC end sequences of Library PTB
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 850)
AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:chimpbes@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT Clones are derived from the chimpanzee BAC library PTB This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.
PRIMERS
Sequencing: -21M13
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI.
FEATURES Location/Qualifiers
source 1. .850
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-149N08.F"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"
ORIGIN

Query Match 56.9%; Score 22.2; DB 10; Length 850;
Best Local Similarity 77.1%; Pred. No. 1.2e+02;
Matches 27; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```
Qy          3 AGCGAGCGTCTAATTCTAGAACACTCACACGCGTC 37
              | | ||| || ||| | | |||||||||||||| ||
Db        627 ACCCAGCCTCAAATACCAAAACACTCACACGCCTC 661
```

RESULT 7
AI179396/c
LOCUS AI179396 506 bp mRNA linear EST 08-JAN-1999
DEFINITION EST223098 Normalized rat spleen, Bento Soares Rattus sp. cDNA clone RSPCG66 3' end, mRNA sequence.
ACCESSION AI179396
VERSION AI179396.1 GI:4135155
KEYWORDS EST.
SOURCE Rattus sp.
ORGANISM Rattus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 506)
AUTHORS Lee,N.H., Glodek,A., Chandra,I., Mason,T.M., Quackenbush,J., Kerlavage,A.R. and Adams,M.D.
TITLE Rat Genome Project: Generation of a Rat EST (REST) Catalog & Rat Gene Index
JOURNAL Unpublished (1998)
COMMENT On Oct 8, 1998 this sequence version replaced gi:3730034.
Contact: Lee, NH
The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-3529
Fax: (301)-838-0208
Email: nhlee@tigr.org
Seq primer: M13-21.
FEATURES Location/Qualifiers
source 1. .506
/organism="Rattus sp."
/mol_type="mRNA"
/db_xref="taxon:10118"
/clone="RSPCG66"
/clone_lib="Normalized rat spleen, Bento Soares"
/note="Organ: spleen; Vector: pT7T3Pac; Site_1: EcoRI; Site_2: NotI"
ORIGIN

Query Match 56.4%; Score 22; DB 1; Length 506;
Best Local Similarity 73.7%; Pred. No. 1.4e+02;
Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```
Qy          2 TAGCGAGCGTCTAATTCTAGAACACTCACACGCGTCTT 39
              ||| || |||| ||  | ||||||||||    ||||||
Db        506 TAGGGACCGTCCAACCCGAGAACACTCAACTTCGTCTT 469
```

RESULT 8
AQ494887
LOCUS AQ494887 439 bp DNA linear GSS 28-APR-1999
DEFINITION HS_5195_A1_C02_SP6E RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=771 Col=3 Row=E, genomic survey sequence.
ACCESSION AQ494887
VERSION AQ494887.1 GI:4695054
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 439)
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
PUBMED 10449764
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm) or from Resear h Genetics (info@resgen.com). BAC end Web Server: http://www.htsc.washington.edu
Plate: 771 row: E column: 3
Seq primer: SP6
Class: BAC ends
High quality sequence stop: 439.
FEATURES Location/Qualifiers
source 1. .439
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=771 Col=3 Row=E"
/sex="male"
/clone_lib="RPCI-11 Human Male BAC Library"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACe3.6 vector at EcoRI sites"
ORIGIN

Query Match 55.9%; Score 21.8; DB 9; Length 439;
Best Local Similarity 78.8%; Pred. No. 1.7e+02;
Matches 26; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
Qy          7 AGCGTCTAATTCTAGAACACTCACACGCGTCTT 39
              | | |||| |||||||||| ||||| | || |
Db        374 ACCCTCTACTTCTAGAACACACACACCCATCAT 406
```

RESULT 9
CE120785
LOCUS CE120785 635 bp DNA linear GSS 25-SEP-2003
DEFINITION tigr-gss-dog-17000325906449 Dog Library Canis familiaris genomic, genomic survey sequence.
ACCESSION CE120785
VERSION CE120785.1 GI:35216186
KEYWORDS GSS.
SOURCE Canis familiaris (dog)
ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE 1 (bases 1 to 635)
AUTHORS Kirkness,E.F., Bafna,V., Halpern,A.L., Levy,S., Remington,K., Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and Venter,J.C.
TITLE The dog genome: survey sequencing and comparative analysis
JOURNAL Science 301 (5641), 1898-1903 (2003)
PUBMED 14512627
COMMENT Contact: Kirkness EF
The Institute for Genomic Research
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirknes@tigr.org
Class: shotgun.
FEATURES Location/Qualifiers
source 1. .635
/organism="Canis familiaris"

```
                     /mol_type="genomic DNA"
                     /strain="Standard Poodle"
                     /db_xref="taxon:9615"
                     /clone_lib="Dog Library"
                     /note="Site_1: BstXI; Libraries were prepared from
                     peripheral blood"
ORIGIN

  Query Match             55.9%;  Score 21.8;  DB 9;  Length 635;
  Best Local Similarity   92.0%;  Pred. No. 1.7e+02;
  Matches   23;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         10 GTCTAATTCTAGAACACTCACACGC 34
              ||| ||||||||||||||||||| |
Db        195 GTCCAATTCTAGAACACTCACACAC 219


RESULT 10
CO731803/c
LOCUS       CO731803                935 bp    mRNA    linear   EST 29-DEC-2004
DEFINITION  JBW027A05.b_047.abi Pineapple green mature fruit cDNA library
            Ananas comosus cDNA clone JBW027A05 similar to 27k
            vesicle-associated membrane protein-associated protein, mRNA
            sequence.
ACCESSION   CO731803
VERSION     CO731803.1  GI:56848975
KEYWORDS    EST.
SOURCE      Ananas comosus (pineapple)
  ORGANISM  Ananas comosus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Bromeliaceae;
            Ananas.
REFERENCE   1  (bases 1 to 935)
  AUTHORS   Moyle,R., Fairbairn,D.J., Ripi,J., Crowe,M. and Botella,J.R.
  TITLE     Developing pineapple fruit has a small transcriptome dominated by
            metallothionein
  JOURNAL   J. Exp. Bot. 56 (409), 101-112 (2005)
   PUBMED   15520025
COMMENT     Contact: Crowe ML
            School of Life Sciences
            The University of Queensland
            Brisbane, Queensland 4072, Australia
            Tel: +617 33462621
            Email: m.crowe@imb.uq.edu.au
            Insert Length: 935   Std Error: 0.00.
FEATURES             Location/Qualifiers
     source          1. .935
                     /organism="Ananas comosus"
                     /mol_type="mRNA"
                     /cultivar="smooth cayenne"
                     /db_xref="taxon:4615"
                     /clone="JBW027A05"
                     /tissue_type="Fruit flesh"
                     /dev_stage="Mature green fruit"
                     /clone_lib="Pineapple green mature fruit cDNA library"
ORIGIN

  Query Match             55.9%;  Score 21.8;  DB 7;  Length 935;
  Best Local Similarity   78.8%;  Pred. No. 1.8e+02;
  Matches   26;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          2 TAGCGAGCGTCTAATTCTAGAACACTCACACGC 34
              ||| | |  |||||||||| ||||| ||||| |
Db        897 TAGTGTGTTTCTAATTCTACAACACACACACAC 865


RESULT 11
CL491525
LOCUS       CL491525                1032 bp    DNA     linear   GSS 01-APR-2004
DEFINITION  SAIL_557_D07.v1 SAIL Collection Arabidopsis thaliana genomic clone
            SAIL_557_D07.v1, genomic survey sequence.
ACCESSION   CL491525
VERSION     CL491525.1  GI:45976434
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1032)
  AUTHORS   Sessions,A., Burke,E., Presting,G., Aux,G., McElver,J., Patton,D.,
            Dietrich,B., Ho,P., Bacwaden,J., Ko,C., Clarke,J.D., Cotton,D.,
            Bullis,D., Snell,J., Miguel,T., Hutchison,D., Kimmerly,B.,
            Mitzel,T., Katagiri,F., Glazebrook,J., Law,M. and Goff,S.A.
  TITLE     A high-throughput Arabidopsis reverse genetics system
  JOURNAL   Plant Cell 14 (12), 2985-2994 (2002)
   PUBMED   12468722
COMMENT     Contact: Sessions A
            Applied Trait Genetics
            Syngenta Biotechnology Inc.
            3054 Cornwallis Rd., Research Triangle Park, NC 27709, USA
            Email: allen.sessions@syngenta.com
            ABRC Stock Number CS823589; T-DNA left border flanking sequences of
            Syngenta Arabidopsis Insertion Library (SAIL) lines are available
            through the Arabidopsis Biological Resource Center (ABRC).
            Sequences represent a pool of amplified genomic regions and not
            single contiguous sequences.
            Class: TDNA tagged.
FEATURES             Location/Qualifiers
     source          1. .1032
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /ecotype="Columbia"
                     /db_xref="taxon:3702"
                     /clone="SAIL_557_D07.v1"
                     /clone_lib="SAIL Collection"
                     /note="T-DNA left border sequences were isolated using a
                     modified TAIL-PCR strategy"
ORIGIN

  Query Match             55.9%;  Score 21.8;  DB 10;  Length 1032;
  Best Local Similarity   78.8%;  Pred. No. 1.8e+02;
  Matches   26;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          5 CGAGCGTCTAATTCTAGAACACTCACACGCGTC 37
              || |||||||||| || ||| | |||  ||||
Db         90 CGGGCGTCTAATTCCAGTACATTAACAAACGTC 122


RESULT 12
AW407061
LOCUS       AW407061                295 bp    mRNA    linear   EST 16-FEB-2000
DEFINITION  UI-HF-BL0-acy-b-01-0-UI.r1 NIH_MGC_37 Homo sapiens cDNA clone
            IMAGE:3060433 5', mRNA sequence.
ACCESSION   AW407061
VERSION     AW407061.1  GI:6926118
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 295)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Eco RI site shown at the beginning of the sequence.
             Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
             cDNA Library Preparation: M.B. Soares Lab
             cDNA Library Arrayed by: M.B. Soares Lab
             DNA Sequencing by: M.B. Soares Lab
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
```

```
          Seq primer: M13 Forward.
FEATURES             Location/Qualifiers
     source          1. .295
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3060433"
                     /tissue_type="lymph"
                     /cell_type="germinal center B cells"
                     /cell_line="MGC85"
                     /lab_host="DH10B (LTI)"
                     /clone_lib="NIH_MGC_37"
                     /note="Vector: pT7T3-Pac; Site_1: NotI; Site_2: Eco RI;
                     Constructed from size fractionated cytoplasmic mRNA
                     (1.5-2.5kb). Directionally cloned. Cells provided by Louis
                     M. Staudt, Ph.D. Library preparation by Maria de Fatima
                     Bonaldo, Ph.D. and M. Bento Soares, Ph.D."
ORIGIN

  Query Match             54.9%;  Score 21.4;  DB 1;  Length 295;
  Best Local Similarity   71.8%;  Pred. No. 2.4e+02;
  Matches   28;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          1 ATAGCGAGCGTCTAATTCTAGAACACTCACACGCGTCTT 39
              |||| || |||| ||| | ||||||  ||  | | ||||
Db        254 ATAGGGATCGTCCAATCCGAGAACATACAACCTCCTCTT 292


RESULT 13
BM851087
LOCUS       BM851087                325 bp    mRNA    linear   EST 06-MAR-2002
DEFINITION  K-EST0131832 S19N665307 Homo sapiens cDNA clone S19N665307-18-H10
            5', mRNA sequence.
ACCESSION   BM851087
VERSION     BM851087.1  GI:19207486
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 325)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 18  row: H  column: 10
            High quality sequence stop: 325.
FEATURES             Location/Qualifiers
     source          1. .325
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="S19N665307-18-H10"
                     /sex="M"
                     /lab_host="Top10F'"
                     /clone_lib="S19N665307"
                     /note="Organ: Stomach; Vector: pCNS; Site_1: EcoRI;
                     Site_2: NotI; The poly (A)+ RNA was dephosphorylated with
                     bacterial alkaline phosphatase (BAP) and then decapped
                     with tabacco acid pyrophosphatase (TAP). The decapped
                     intact mRNA was ligated with DNA-RNA linker including EcoR
                     I site by treatment of T4 RNA ligase and the first strand
                     cDNA was synthesized from oligo dT-selected mRNA by
                     priming with dT-tailed vector. The dT-tailed vector was
                     adjusted to have about 60nt. The cDNA vector was
                     circularized with E. coli DNA ligase after digestion of
                     EcoRI which site is also included in vector. An RNA strand
                     converted to a DNA strand by Okayama-Berg method. The
                     obtained cDNA vectors were used for transformation of
                     competent cells E. coli Top10F' by electroporation method.
                     The cDNA libraries constructed by this method are
                     full-length enriched cDNA library."
ORIGIN

  Query Match             54.9%;  Score 21.4;  DB 3;  Length 325;
  Best Local Similarity   71.8%;  Pred. No. 2.4e+02;
  Matches   28;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          1 ATAGCGAGCGTCTAATTCTAGAACACTCACACGCGTCTT 39
              |||| || |||| ||  | ||||||  || || | ||||
Db        287 ATAGGGATCGTCCAACCCGAGAACATACAAACTCCTCTT 325


RESULT 14
BM851189
LOCUS       BM851189                445 bp    mRNA    linear   EST 06-MAR-2002
DEFINITION  K-EST0131956 S19N665307 Homo sapiens cDNA clone S19N665307-17-C02
            5', mRNA sequence.
ACCESSION   BM851189
VERSION     BM851189.1  GI:19207588
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 445)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 17  row: C  column: 02
            High quality sequence stop: 445.
FEATURES             Location/Qualifiers
     source          1. .445
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="S19N665307-17-C02"
                     /sex="M"
                     /lab_host="Top10F'"
                     /clone_lib="S19N665307"
                     /note="Organ: Stomach; Vector: pCNS; Site_1: EcoRI;
                     Site_2: NotI; The poly (A)+ RNA was dephosphorylated with
                     bacterial alkaline phosphatase (BAP) and then decapped
                     with tabacco acid pyrophosphatase (TAP). The decapped
                     intact mRNA was ligated with DNA-RNA linker including EcoR
                     I site by treatment of T4 RNA ligase and the first strand
                     cDNA was synthesized from oligo dT-selected mRNA by
                     priming with dT-tailed vector. The dT-tailed vector was
                     adjusted to have about 60nt. The cDNA vector was
                     circularized with E. coli DNA ligase after digestion of
                     EcoRI which site is also included in vector. An RNA strand
                     converted to a DNA strand by Okayama-Berg method. The
                     obtained cDNA vectors were used for transformation of
                     competent cells E. coli Top10F' by electroporation method.
                     The cDNA libraries constructed by this method are
                     full-length enriched cDNA library."
ORIGIN
```

```
  Query Match             54.9%;  Score 21.4;  DB 3;  Length 445;
  Best Local Similarity   71.8%;  Pred. No. 2.5e+02;
  Matches   28;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          1 ATAGCGAGCGTCTAATTCTAGAACACTCACACGCGTCTT 39
              |||| || |||| ||  | ||||||  || || | ||||
Db        287 ATAGGGATCGTCCAACCCGAGAACATACAAACTCCTCTT 325


RESULT 15
AA695264/c
LOCUS       AA695264                 504 bp    mRNA    linear   EST 23-APR-2001
DEFINITION  GM02505.5prime GM Drosophila melanogaster ovary BlueScript
            Drosophila melanogaster cDNA clone GM02505 5, mRNA sequence.
ACCESSION   AA695264
VERSION     AA695264.2  GI:13760555
KEYWORDS    EST.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 504)
  AUTHORS   Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
            Lewis,S. and Rubin,G.M.
  TITLE     BDGP/HHMI Drosophila EST Project
  JOURNAL   Unpublished (2001)
COMMENT     On Dec 18, 1997 this sequence version replaced gi:2697884.
            Other_ESTs: GM02505.3prime
            Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            hit genomic AE003504: arm:X [16678765,16977278]
            estimated-cyto:15B4-15E7: 04/10/2001
            Plate: GM.25  row: A  column: 5
            High quality sequence stop: 459
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .504
                     /organism="Drosophila melanogaster"
                     /mol_type="mRNA"
                     /db_xref="taxon:7227"
                     /clone="GM02505"
                     /sex="female"
                     /dev_stage="newly eclosed females: germarium-stage 6"
                     /lab_host="SOLR"
                     /clone_lib="GM Drosophila melanogaster ovary BlueScript"
                     /note="Organ: ovary; Vector: BlueScript SK; Site_1: EcoRI;
                     Site_2: XhoI; Constructed using Stratagene ZAP-cDNA
                     Synthesis kit. Oligo dT-primed and directionally cloned at
                     EcoRI and XhoI in BlueScript SK(+/-)"
ORIGIN

  Query Match             54.9%;  Score 21.4;  DB 1;  Length 504;
  Best Local Similarity   71.8%;  Pred. No. 2.5e+02;
  Matches   28;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          1 ATAGCGAGCGTCTAATTCTAGAACACTCACACGCGTCTT 39
              ||| ||||||| ||  | |||||   | |||||| | ||
Db        206 ATAACGAGCGTATATATATAGAAACATAACACGCATATT 168


Search completed: January 30, 2006, 03:26:20
Job time : 1560 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: January 30, 2006, 01:49:51 ; Search time 73 Seconds
(without alignments)
949.656 Million cell updates/sec

Title: US-10-663-033-REVCOMP4-SEQ3
Perfect score: 39
Sequence: 1 atagcgagcgtctaattctagaacactcacacgcgtctt 39

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 0.1

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/1_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP_COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE_COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		21.4	54.9	2182	3	US-09-270-767-13447	Sequence 13447, A
2	c	20.4	52.3	14641	3	US-09-949-016-15834	Sequence 15834, A
3	c	20	51.3	20	3	US-09-433-598-4	Sequence 4, Appli
4	c	20	51.3	20	3	US-09-848-838A-4	Sequence 4, Appli
5		20	51.3	1119	3	US-09-433-598-1	Sequence 1, Appli
6		20	51.3	1119	3	US-09-848-838A-1	Sequence 1, Appli
7		19.8	50.8	386	3	US-09-513-999C-1023	Sequence 1023, Ap
8		19.8	50.8	554	3	US-09-621-976-1260	Sequence 1260, Ap
9		19.6	50.3	1807	3	US-09-902-540-316	Sequence 316, App
10		19	48.7	19	3	US-09-433-598-3	Sequence 3, Appli
11		19	48.7	19	3	US-09-848-838A-3	Sequence 3, Appli
12	c	19	48.7	4515	3	US-09-328-352-2920	Sequence 2920, Ap
13		19	48.7	8874	3	US-08-894-344C-1	Sequence 1, Appli
14		19	48.7	8874	3	US-09-678-023A-1	Sequence 1, Appli
15		19	48.7	9621	3	US-09-125-028-1	Sequence 1, Appli
16	c	19	48.7	54161	3	US-09-949-016-11905	Sequence 11905, A
17	c	18.8	48.2	92581	3	US-09-949-016-12182	Sequence 12182, A
18	c	18.8	48.2	92581	3	US-09-949-016-16542	Sequence 16542, A
19		18.6	47.7	519	3	US-09-248-796A-2238	Sequence 2238, Ap
20		18.6	47.7	930	4	US-09-605-703B-2563	Sequence 2563, Ap
21	c	18.6	47.7	1215	3	US-09-248-796A-2972	Sequence 2972, Ap
22		18.6	47.7	2502	2	US-08-062-472B-2	Sequence 2, Appli
23	c	18.6	47.7	17341	3	US-09-415-946-1	Sequence 1, Appli
24		18.6	47.7	32042	3	US-09-245-281-44	Sequence 44, Appl
25		18.6	47.7	32042	3	US-09-340-620A-63	Sequence 63, Appl
26		18.6	47.7	32042	3	US-09-728-721-63	Sequence 63, Appl
27		18.6	47.7	39553	3	US-09-949-002-810	Sequence 810, App
28		18.6	47.7	58111	3	US-09-949-002-673	Sequence 673, App
29		18.6	47.7	109925	3	US-09-949-016-13210	Sequence 13210, A
30	c	18.4	47.2	2190	2	US-09-036-582-37	Sequence 37, Appl
31	c	18.4	47.2	2190	3	US-09-318-141-37	Sequence 37, Appl
32	c	18.4	47.2	2819	3	US-10-104-047-264	Sequence 264, App
33		18.4	47.2	39690	3	US-09-949-016-15079	Sequence 15079, A
34	c	18.4	47.2	52199	3	US-09-949-016-12763	Sequence 12763, A
35	c	18.4	47.2	52203	3	US-09-949-016-16330	Sequence 16330, A
36	c	18.4	47.2	70559	3	US-09-409-800B-1	Sequence 1, Appli
37		18.4	47.2	4403765	3	US-09-103-840A-2	Sequence 2, Appli
38		18.4	47.2	4411529	3	US-09-103-840A-1	Sequence 1, Appli
39	c	18.2	46.7	490	3	US-09-621-976-294	Sequence 294, App
40	c	18.2	46.7	601	3	US-09-949-016-65208	Sequence 65208, A
41		18.2	46.7	601	3	US-09-949-016-83545	Sequence 83545, A
42		18.2	46.7	601	3	US-09-949-016-83546	Sequence 83546, A
43		18.2	46.7	601	3	US-09-949-016-83547	Sequence 83547, A
44		18.2	46.7	601	3	US-09-949-016-83548	Sequence 83548, A
45	c	18.2	46.7	653	3	US-09-533-559-6394	Sequence 6394, Ap

ALIGNMENTS

RESULT 1
US-09-270-767-13447
; Sequence 13447, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13447
; LENGTH: 2182
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-13447

Query Match 54.9%; Score 21.4; DB 3; Length 2182;
Best Local Similarity 71.8%; Pred. No. 7.1;
Matches 28; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

```
Qy          1 ATAGCGAGCGTCTAATTCTAGAACACTCACACGCGTCTT 39
              ||| ||||||| ||  | |||||  | |||||| | ||
Db       1978 ATAACGAGCGTATATATATAGAAACATAACACGCATATT 2016
```

RESULT 2
US-09-949-016-15834/c
; Sequence 15834, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15834

```
;   LENGTH: 14641
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(14641)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15834

  Query Match            52.3%;  Score 20.4;  DB 3;  Length 14641;
  Best Local Similarity  71.1%;  Pred. No. 33;
  Matches   27;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          1 ATAGCGAGCGTCTAATTCTAGAACACTCACACGCGTCT 38
              |||| ||| || |||| |||| | ||| | | ||  ||
Db       5467 ATAGAGAGTGTTTAATGCTAGGAGACTAAGAGGCTACT 5430


RESULT 3
US-09-433-598-4/c
; Sequence 4, Application US/09433598
; Patent No. 6238894
; GENERAL INFORMATION:
;  APPLICANT: Taylor Dr., Diane
;  APPLICANT: Wang, Ge
;  APPLICANT: Palcic, Monica
;  TITLE OF INVENTION: Alpha1, 2 Fucosyltransferase
;  FILE REFERENCE: 07254/061001
;  CURRENT APPLICATION NUMBER: US/09/433,598
;  CURRENT FILING DATE: 1998-11-02
;  EARLIER APPLICATION NUMBER: 60/107268
;  EARLIER FILING DATE: 1998-11-04
;  NUMBER OF SEQ ID NOS: 23
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
;   LENGTH: 20
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:
;   OTHER INFORMATION: oligonucleotide
US-09-433-598-4

  Query Match            51.3%;  Score 20;  DB 3;  Length 20;
  Best Local Similarity  100.0%;  Pred. No. 9.6;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATAGCGAGCGTCTAATTCTA 20
              ||||||||||||||||||||
Db         20 ATAGCGAGCGTCTAATTCTA 1


RESULT 4
US-09-848-838A-4/c
; Sequence 4, Application US/09848838A
; Patent No. 6670160
; GENERAL INFORMATION:
;  APPLICANT: Taylor, Diane
;  APPLICANT:  Wang, Ge
;  APPLICANT:  Palcic, Monica
;  TITLE OF INVENTION: ALPHA 1,2-FUCOSYLTRANSFERASE
;  FILE REFERENCE: 07254-061002
;  CURRENT APPLICATION NUMBER: US/09/848,838A
;  CURRENT FILING DATE:  2002-05-21
;  PRIOR APPLICATION NUMBER: US 09/433,598
;  PRIOR FILING DATE: 1999-11-02
;  PRIOR APPLICATION NUMBER: US 60/107,268
;  PRIOR FILING DATE: 1998-11-04
;  NUMBER OF SEQ ID NOS: 23
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;   LENGTH: 20
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: synthetically generated oligonucleotide
US-09-848-838A-4

  Query Match            51.3%;  Score 20;  DB 3;  Length 20;
  Best Local Similarity  100.0%;  Pred. No. 9.6;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATAGCGAGCGTCTAATTCTA 20
              ||||||||||||||||||||
Db         20 ATAGCGAGCGTCTAATTCTA 1


RESULT 5
US-09-433-598-1
; Sequence 1, Application US/09433598
; Patent No. 6238894
; GENERAL INFORMATION:
;  APPLICANT: Taylor Dr., Diane
;  APPLICANT: Wang, Ge
;  APPLICANT: Palcic, Monica
;  TITLE OF INVENTION: Alpha1, 2 Fucosyltransferase
;  FILE REFERENCE: 07254/061001
;  CURRENT APPLICATION NUMBER: US/09/433,598
;  CURRENT FILING DATE: 1998-11-02
;  EARLIER APPLICATION NUMBER: 60/107268
;  EARLIER FILING DATE: 1998-11-04
;  NUMBER OF SEQ ID NOS: 23
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 1119
;   TYPE: DNA
;   ORGANISM: Helicobacter pylori
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (137)..(1036)
US-09-433-598-1

  Query Match            51.3%;  Score 20;  DB 3;  Length 1119;
  Best Local Similarity  100.0%;  Pred. No. 26;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATAGCGAGCGTCTAATTCTA 20
              ||||||||||||||||||||
Db       1100 ATAGCGAGCGTCTAATTCTA 1119


RESULT 6
US-09-848-838A-1
; Sequence 1, Application US/09848838A
; Patent No. 6670160
; GENERAL INFORMATION:
;  APPLICANT: Taylor, Diane
;  APPLICANT:  Wang, Ge
;  APPLICANT:  Palcic, Monica
;  TITLE OF INVENTION: ALPHA 1,2-FUCOSYLTRANSFERASE
;  FILE REFERENCE: 07254-061002
;  CURRENT APPLICATION NUMBER: US/09/848,838A
;  CURRENT FILING DATE:  2002-05-21
;  PRIOR APPLICATION NUMBER: US 09/433,598
;  PRIOR FILING DATE: 1999-11-02
;  PRIOR APPLICATION NUMBER: US 60/107,268
;  PRIOR FILING DATE: 1998-11-04
;  NUMBER OF SEQ ID NOS: 23
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 1119
;   TYPE: DNA
;   ORGANISM: Helicobacter pylori
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (137)...(1036)
```

US-09-848-838A-1

Query Match 51.3%; Score 20; DB 3; Length 1119;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATAGCGAGCGTCTAATTCTA 20
              ||||||||||||||||||||
Db       1100 ATAGCGAGCGTCTAATTCTA 1119
```

RESULT 7
US-09-513-999C-1023
; Sequence 1023, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 1023
; LENGTH: 386
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 87..386
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 383
; OTHER INFORMATION: y=c or t
US-09-513-999C-1023

Query Match 50.8%; Score 19.8; DB 3; Length 386;
Best Local Similarity 69.2%; Pred. No. 25;
Matches 27; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

```
Qy          1 ATAGCGAGCGTCTAATTCTAGAACACTCACACGCGTCTT 39
              |||| || |||| ||  | ||||||  ||  | | ||||
Db        310 ATAGGGATCGTCCAACCCGAGAACATACAACCTCCTCTT 348
```

RESULT 8
US-09-621-976-1260
; Sequence 1260, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 1260
; LENGTH: 554
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 121..552
; NAME/KEY: misc_feature
; LOCATION: 32
; OTHER INFORMATION: n=a, g, c or t
US-09-621-976-1260

Query Match 50.8%; Score 19.8; DB 3; Length 554;
Best Local Similarity 69.2%; Pred. No. 27;
Matches 27; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

```
Qy          1 ATAGCGAGCGTCTAATTCTAGAACACTCACACGCGTCTT 39
              |||| || |||| ||  | ||||||  ||  | | ||||
Db        344 ATAGGGATCGTCCAACCCGAGAACATACAACCTCCTCTT 382
```

RESULT 9
US-09-902-540-316
; Sequence 316, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 316
; LENGTH: 1807
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-316

Query Match 50.3%; Score 19.6; DB 3; Length 1807;
Best Local Similarity 73.5%; Pred. No. 45;
Matches 25; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```
Qy          4 GCGAGCGTCTAATTCTAGAACACTCACACGCGTC 37
              |||  ||||  ||||||||||   |||||  |||
Db       1261 GCGCCCGTCGCATTCTAGAACGAGCACACCGGTC 1294
```

RESULT 10
US-09-433-598-3
; Sequence 3, Application US/09433598
; Patent No. 6238894
; GENERAL INFORMATION:
; APPLICANT: Taylor Dr., Diane
; APPLICANT: Wang, Ge
; APPLICANT: Palcic, Monica
; TITLE OF INVENTION: Alpha1, 2 Fucosyltransferase
; FILE REFERENCE: 07254/061001
; CURRENT APPLICATION NUMBER: US/09/433,598
; CURRENT FILING DATE: 1998-11-02
; EARLIER APPLICATION NUMBER: 60/107268
; EARLIER FILING DATE: 1998-11-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 19
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: oligonucleotide
US-09-433-598-3

Query Match 48.7%; Score 19; DB 3; Length 19;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         21 GAACACTCACACGCGTCTT 39
```

```
              ||||||||||||||||||
Db          1 GAACACTCACACGCGTCTT 19


RESULT 11
US-09-848-838A-3
; Sequence 3, Application US/09848838A
; Patent No. 6670160
; GENERAL INFORMATION:
;  APPLICANT: Taylor, Diane
;  APPLICANT:  Wang, Ge
;  APPLICANT:  Palcic, Monica
;  TITLE OF INVENTION: ALPHA 1,2-FUCOSYLTRANSFERASE
;  FILE REFERENCE: 07254-061002
;  CURRENT APPLICATION NUMBER: US/09/848,838A
;  CURRENT FILING DATE:  2002-05-21
;  PRIOR APPLICATION NUMBER: US 09/433,598
;  PRIOR FILING DATE: 1999-11-02
;  PRIOR APPLICATION NUMBER: US 60/107,268
;  PRIOR FILING DATE: 1998-11-04
;  NUMBER OF SEQ ID NOS: 23
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 19
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: synthetically generated oligonucleotide
US-09-848-838A-3

  Query Match             48.7%;  Score 19;  DB 3;  Length 19;
  Best Local Similarity   100.0%;  Pred. No. 27;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         21 GAACACTCACACGCGTCTT 39
              ||||||||||||||||||
Db          1 GAACACTCACACGCGTCTT 19


RESULT 12
US-09-328-352-2920/c
; Sequence 2920, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
;  APPLICANT: Gary L. Breton et al.
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
;  TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: GTC99-03PA
;  CURRENT APPLICATION NUMBER: US/09/328,352
;  CURRENT FILING DATE: 1999-06-04
;  NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 2920
;   LENGTH: 4515
;   TYPE: DNA
;   ORGANISM: Acinetobacter baumannii
US-09-328-352-2920

  Query Match             48.7%;  Score 19;  DB 3;  Length 4515;
  Best Local Similarity   81.5%;  Pred. No. 1.1e+02;
  Matches   22;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          6 GAGCGTCTAATTCTAGAACACTCACAC 32
              |||| |||||| | ||| |||||| ||
Db       4321 GAGCATCTAATGCAAGACCACTCAAAC 4295


RESULT 13
US-08-894-344C-1
; Sequence 1, Application US/08894344C
; Patent No. 6172196
;  GENERAL INFORMATION:
;    APPLICANT:  KAWASAKI, Hideki
;    APPLICANT:  TOKAI, Masaya
;    APPLICANT:  KIKUCHI, Yasuhiro
;    APPLICANT:  OUCHI, Kozo
;    TITLE OF INVENTION:  DNA ENCODING PROTEIN COMPLEMENTING
;    TITLE OF INVENTION:  YEAST
;    TITLE OF INVENTION:  LOW TEMPERATURE-SENSITIVE FERMENTABILITY
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FITZPATRICK, CELLA, HARPER & SCINTO
;      STREET:  30 Rockefeller Plaza
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10112-3801
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette - 3.50 inch, 1440 Kb storage.
;      COMPUTER:  IBM PS/V
;      OPERATING SYSTEM:  MS-DOS Ver3.30
;      SOFTWARE:  PATENT AID Ver1.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/894,344C
;      FILING DATE:  15-AUGUST-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP343700/95
;      FILING DATE:  28-DECEMBER-1995
;      APPLICATION NUMBER:  PCT/JP96/03862
;      FILING DATE:  27-DECEMBER-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Perry, Lawrence S.
;      REGISTRATION NUMBER:  31865
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-218-2100
;      TELEFAX:  212-218-2200
;  INFORMATION FOR SEQ ID NO:  1 :
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  8874 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Genomic DNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Saccharomyces cerevisiae
;      STRAIN:  X2180-1B
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1 to 8874
;      IDENTIFICATION METHOD:  E
;    FEATURE:
;      NAME/KEY:  cleavage-site
;      LOCATION:  1291 to 1296
;      IDENTIFICATION METHOD:  S
;    FEATURE:
;      NAME/KEY:  cleavage-site
;      LOCATION:  4388 to 4393
;      IDENTIFICATION METHOD:  S
;    FEATURE:
;      NAME/KEY:  cleavage-site
;      LOCATION:  5927 to 5032
;      IDENTIFICATION METHOD:  S
;    FEATURE:
;      NAME/KEY:  cleavage-site
;      LOCATION:  7675 to 7680
;      IDENTIFICATION METHOD:  S
US-08-894-344C-1

  Query Match             48.7%;  Score 19;  DB 3;  Length 8874;
  Best Local Similarity   81.5%;  Pred. No. 1.3e+02;
  Matches   22;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         11 TCTAATTCTAGAACACTCACACGCGTC 37
              | ||| |||||||||| || | |||||
Db       7215 TTTAACTCTAGAACACACAAAAGCGTC 7241
```

```
RESULT 14
US-09-678-023A-1
; Sequence 1, Application US/09678023A
; Patent No. 6924132
;  GENERAL INFORMATION:
;       APPLICANT: KAWASAKI, Hideki
;                  TOKAI, Masaya
;                  KIKUCHI, Yasuhiro
;                  OUCHI, Kozo
;       TITLE OF INVENTION: DNA ENCODING PROTEIN COMPLEMENTING
;                           YEAST
;                           LOW TEMPERATURE-SENSITIVE FERMENTABILITY
;       NUMBER OF SEQUENCES: 2
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: FITZPATRICK, CELLA, HARPER & SCINTO
;            STREET: 30 Rockefeller Plaza
;            CITY: New York
;            STATE: New York
;            COUNTRY: U.S.A.
;            ZIP: 10112-3801
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Diskette - 3.50 inch, 1440 Kb storage.
;            COMPUTER: IBM PS/V
;            OPERATING SYSTEM: MS-DOS Ver3.30
;            SOFTWARE: PATENT AID Ver1.0
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/678,023A
;            FILING DATE: 04-Oct-2000
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 08/894,344
;            FILING DATE: 15-AUGUST-1997
;            APPLICATION NUMBER: JP343700/95
;            FILING DATE: 28-DECEMBER-1995
;            APPLICATION NUMBER: PCT/JP96/03862
;            FILING DATE: 27-DECEMBER-1996
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Perry, Lawrence S.
;            REGISTRATION NUMBER: 31865
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 212-218-2100
;            TELEFAX: 212-218-2200
;  INFORMATION FOR SEQ ID NO: 1 :
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 8874 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: double
;            TOPOLOGY: linear
;       MOLECULE TYPE: Genomic DNA
;       ORIGINAL SOURCE:
;            ORGANISM: Saccharomyces cerevisiae
;            STRAIN: X2180-1B
;       FEATURE:
;            NAME/KEY:  CDS
;            LOCATION:  1 to 8874
;            IDENTIFICATION METHOD:  E
;       FEATURE:
;            NAME/KEY:  cleavage-site
;            LOCATION:  1291 to 1296
;            IDENTIFICATION METHOD:  S
;       FEATURE:
;            NAME/KEY:  cleavage-site
;            LOCATION:  4388 to 4393
;            IDENTIFICATION METHOD:  S
;       FEATURE:
;            NAME/KEY:  cleavage-site
;            LOCATION:  5027 to 5032
;            IDENTIFICATION METHOD:  S
;       FEATURE:
;            NAME/KEY:  cleavage-site
;            LOCATION:  7675 to 7680
;            IDENTIFICATION METHOD:  S
;       SEQUENCE DESCRIPTION: SEQ ID NO: 1
US-09-678-023A-1

  Query Match             48.7%;  Score 19;  DB 3;  Length 8874;
  Best Local Similarity   81.5%;  Pred. No. 1.3e+02;
  Matches   22;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         11 TCTAATTCTAGAACACTCACACGCGTC 37
              | ||| |||||||||| || | |||||
Db       7215 TTTAACTCTAGAACACACAAAAGCGTC 7241


RESULT 15
US-09-125-028-1
; Sequence 1, Application US/09125028A
; Patent No. 6190707
; GENERAL INFORMATION:
;  APPLICANT: WADOUX, Isabelle
;  APPLICANT: COLAVIZZA, Didier
;  APPLICANT: LOIEZ, Annie
;  TITLE OF INVENTION: NOVEL COLD-SENSITIVE BREAD-MAKING YEASTS
;  FILE REFERENCE: levure sensible froid
;  CURRENT APPLICATION NUMBER: US/09/125,028A
;  CURRENT FILING DATE: 1998-08-07
;  EARLIER APPLICATION NUMBER: PCT/FR97/00254
;  EARLIER FILING DATE: 1997-02-07
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 9621
;   TYPE: DNA
;   ORGANISM: Saccharomyces cerevisiae
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (9318)
;   OTHER INFORMATION: Use of n signifies any of g, a, c or t
US-09-125-028-1

  Query Match             48.7%;  Score 19;  DB 3;  Length 9621;
  Best Local Similarity   81.5%;  Pred. No. 1.3e+02;
  Matches   22;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         11 TCTAATTCTAGAACACTCACACGCGTC 37
              | ||| |||||||||| || | |||||
Db       7590 TTTAACTCTAGAACACACAAAAGCGTC 7616


Search completed: January 30, 2006, 03:28:48
Job time : 75 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: January 30, 2006, 02:34:27 ; Search time 349 Seconds
(without alignments)
924.085 Million cell updates/sec

Title: US-10-663-033-REVCOMP4-SEQ3
Perfect score: 39
Sequence: 1 atagcgagcgtctaattctagaacactcacacgcgtctt 39

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 0.1

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA_Main:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		22.4	57.4	653	3	US-09-988-067B-69	Sequence 69, Appl
2		21	53.8	584	4	US-09-925-065A-336197	Sequence 336197,
3		21	53.8	604	4	US-09-925-065A-336196	Sequence 336196,
c 4		20.6	52.8	2409	6	US-10-369-493-26223	Sequence 26223, A
c 5		20.4	52.3	395	8	US-10-425-115-31553	Sequence 31553, A
c 6		20.4	52.3	488	4	US-09-925-065A-754744	Sequence 754744,
7		20.2	51.8	451	5	US-10-198-846-375	Sequence 375, App
8		20.2	51.8	686	5	US-10-198-846-1034	Sequence 1034, Ap
c 9		20.2	51.8	773	9	US-10-779-543-23653	Sequence 23653, A
c 10		20.2	51.8	2844	5	US-10-098-841-221	Sequence 221, App
c 11		20.2	51.8	2871	8	US-10-788-792-120	Sequence 120, App
c 12		20.2	51.8	2881	5	US-10-087-192-635	Sequence 635, App
c 13		20.2	51.8	2881	7	US-10-058-270A-65	Sequence 65, Appl
c 14		20.2	51.8	3230	5	US-10-198-846-10349	Sequence 10349, A
c 15		20.2	51.8	46404	5	US-10-087-192-634	Sequence 634, App
c 16		20	51.3	20	3	US-09-848-838-4	Sequence 4, Appli
c 17		20	51.3	20	7	US-10-663-033-4	Sequence 4, Appli
c 18		20	51.3	577	4	US-09-925-065A-78749	Sequence 78749, A
c 19		20	51.3	934	3	US-09-895-913A-253	Sequence 253, App
20		20	51.3	1119	3	US-09-848-838-1	Sequence 1, Appli
21		20	51.3	1119	7	US-10-663-033-1	Sequence 1, Appli
22		19.8	50.8	325	7	US-10-242-535A-10789	Sequence 10789, A
23		19.8	50.8	325	7	US-10-085-783A-10789	Sequence 10789, A
c 24		19.8	50.8	325	8	US-10-674-124A-11175	Sequence 11175, A
25		19.8	50.8	370	7	US-10-242-535A-22219	Sequence 22219, A
26		19.8	50.8	370	7	US-10-085-783A-22219	Sequence 22219, A
27		19.8	50.8	438	7	US-10-242-535A-34927	Sequence 34927, A
28		19.8	50.8	438	7	US-10-085-783A-34927	Sequence 34927, A
c 29		19.8	50.8	487	8	US-10-425-115-5839	Sequence 5839, Ap
30		19.8	50.8	488	7	US-10-242-535A-56833	Sequence 56833, A
31		19.8	50.8	488	7	US-10-085-783A-56833	Sequence 56833, A
32		19.8	50.8	569	8	US-10-357-930-35647	Sequence 35647, A
33		19.8	50.8	569	8	US-10-357-930-44454	Sequence 44454, A
34		19.8	50.8	570	8	US-10-357-930-14541	Sequence 14541, A
35		19.8	50.8	731	3	US-09-822-830A-72	Sequence 72, Appl
c 36		19.8	50.8	790	3	US-09-883-152-20	Sequence 20, Appl
c 37		19.8	50.8	800	3	US-09-986-480-32	Sequence 32, Appl
c 38		19.8	50.8	800	9	US-10-863-332-32	Sequence 32, Appl
39		19.8	50.8	826	5	US-10-027-632-145050	Sequence 145050,
40		19.8	50.8	826	5	US-10-027-632-145051	Sequence 145051,
41		19.8	50.8	826	6	US-10-027-632-145050	Sequence 145050,
42		19.8	50.8	826	6	US-10-027-632-145051	Sequence 145051,
43		19.8	50.8	954	6	US-10-264-049-775	Sequence 775, App
44		19.8	50.8	1046	6	US-10-172-118-1642	Sequence 1642, Ap
45		19.8	50.8	1046	7	US-10-342-887-1642	Sequence 1642, Ap

ALIGNMENTS

RESULT 1
US-09-988-067B-69
; Sequence 69, Application US/09988067B
; Publication No. US20030124141A1
; GENERAL INFORMATION:
; APPLICANT: Haas, Rainer
; APPLICANT: Kleanthous, Harold
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Miller, Charles
; APPLICANT: Al-Garawi, Amal
; APPLICANT: Odenbreit, Stefan
; APPLICANT: Meyer, Thomas
; TITLE OF INVENTION: Helicobacter Polypeptides and
; TITLE OF INVENTION: Corresponding Polynucleotide Molecules
; FILE REFERENCE: 06132/040002
; CURRENT APPLICATION NUMBER: US/09/988,067B
; CURRENT FILING DATE: 2003-01-31
; PRIOR APPLICATION NUMBER: US 08/831,309
; PRIOR FILING DATE: 1997-04-01
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 69
; LENGTH: 653
; TYPE: DNA
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (63)...(590)
US-09-988-067B-69

Query Match 57.4%; Score 22.4; DB 3; Length 653;
Best Local Similarity 95.8%; Pred. No. 14;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy         16 TTCTAGAACACTCACACGCGTCTT 39
              | ||||||||||||||||||||||||
Db        547 TGCTAGAACACTCACACGCGTCTT 570
```

RESULT 2
US-09-925-065A-336197
; Sequence 336197, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single

```
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 336197
;   LENGTH: 584
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-336197

  Query Match             53.8%;  Score 21;  DB 4;  Length 584;
  Best Local Similarity   73.0%;  Pred. No. 56;
  Matches   27;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          1 ATAGCGAGCGTCTAATTCTAGAACACTCACACGCGTC 37
              |||    | || | ||||||||| ||||||| || ||
Db        545 ATATTTTGAGTTTTATTCTAGAAAACTCACAAGCCTC 581


RESULT 3
US-09-925-065A-336196
; Sequence 336196, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 336196
;   LENGTH: 604
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-336196

  Query Match             53.8%;  Score 21;  DB 4;  Length 604;
  Best Local Similarity   73.0%;  Pred. No. 56;
  Matches   27;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          1 ATAGCGAGCGTCTAATTCTAGAACACTCACACGCGTC 37
              |||    | || | ||||||||| ||||||| || ||
Db        469 ATATTTTGAGTTTTATTCTAGAAAACTCACAAGCCTC 505


RESULT 4
US-10-369-493-26223/c
; Sequence 26223, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
;  APPLICANT: Cao, Yongwei
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Goldman, Barry S.
;  APPLICANT:  Chen, Xianfeng
;  TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
;  TITLE OF INVENTION:  PLANTS WITH IMPROVED PROPERTIES
;  FILE REFERENCE: 38-10(52052)B
;  CURRENT APPLICATION NUMBER: US/10/369,493
;  CURRENT FILING DATE:  2003-02-28
;  PRIOR APPLICATION NUMBER: US 60/360,039
;  PRIOR FILING DATE: 2002-02-21
;  NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 26223
;   LENGTH: 2409
;   TYPE: DNA
;   ORGANISM: Schizosaccharomyces pombe
US-10-369-493-26223

  Query Match             52.8%;  Score 20.6;  DB 6;  Length 2409;
  Best Local Similarity   74.3%;  Pred. No. 1.2e+02;
  Matches   26;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          4 GCGAGCGTCTAATTCTAGAACACTCACACGCGTCT 38
              | ||||| |||| ||||| ||| | |||||  | |
Db        534 GAGAGCGCCTAAATCTAGGACAATAACACGTTTAT 500


RESULT 5
US-10-425-115-31553/c
; Sequence 31553, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(53222)B
;  CURRENT APPLICATION NUMBER: US/10/425,115
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 31553
;   LENGTH: 395
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(395)
;   OTHER INFORMATION: unsure at all n locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: MRT4577_128786C.1
US-10-425-115-31553

  Query Match             52.3%;  Score 20.4;  DB 8;  Length 395;
  Best Local Similarity   71.1%;  Pred. No. 95;
  Matches   27;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          1 ATAGCGAGCGTCTAATTCTAGAACACTCACACGCGTCT 38
              ||||| ||||||| |||||| | |||  |    | |||
Db        230 ATAGCCAGCGTCTCATTCTACAGCACATAGTTTCTTCT 193


RESULT 6
US-09-925-065A-754744/c
; Sequence 754744, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
```

```
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 754744
;   LENGTH: 488
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-754744

  Query Match             52.3%;  Score 20.4;  DB 4;  Length 488;
  Best Local Similarity   71.1%;  Pred. No. 1e+02;
  Matches   27;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          1 ATAGCGAGCGTCTAATTCTAGAACACTCACACGCGTCT 38
              |||| |||  ||||  ||||  |||| ||||  ||| |
Db        105 ATAGGGAGATTCTATCTCTAACACACACACATCCGTAT 68


RESULT 7
US-10-198-846-375
; Sequence 375, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
;  APPLICANT: Lillie, James
;  APPLICANT:  Xu, Yongyao
;  APPLICANT:  Wang, Youzhen
;  APPLICANT:  Steinmann, Kathleen
;  TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
;  TITLE OF INVENTION:  FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
;  TITLE OF INVENTION:  THERAPY OF BREAST CANCER
;  FILE REFERENCE: MRI-049
;  CURRENT APPLICATION NUMBER: US/10/198,846
;  CURRENT FILING DATE:  2002-07-18
;  PRIOR APPLICATION NUMBER: 60/306,220
;  PRIOR FILING DATE: 2001-07-18
;  NUMBER OF SEQ ID NOS: 14084
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 375
;   LENGTH: 451
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 446
;   OTHER INFORMATION: n = A,T,C or G
US-10-198-846-375

  Query Match             51.8%;  Score 20.2;  DB 5;  Length 451;
  Best Local Similarity   75.8%;  Pred. No. 1.2e+02;
  Matches   25;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          2 TAGCGAGCGTCTAATTCTAGAACACTCACACGC 34
              | | |||| |||| ||||||||||   | ||||
Db        109 TTGGGAGCCTCTATTTCTAGAACAGGAAAACGC 141


RESULT 8
US-10-198-846-1034
; Sequence 1034, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
;  APPLICANT: Lillie, James
;  APPLICANT:  Xu, Yongyao
;  APPLICANT:  Wang, Youzhen
;  APPLICANT:  Steinmann, Kathleen
;  TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
;  TITLE OF INVENTION:  FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
;  TITLE OF INVENTION:  THERAPY OF BREAST CANCER
;  FILE REFERENCE: MRI-049
;  CURRENT APPLICATION NUMBER: US/10/198,846
;  CURRENT FILING DATE:  2002-07-18
;  PRIOR APPLICATION NUMBER: 60/306,220
;  PRIOR FILING DATE: 2001-07-18
;  NUMBER OF SEQ ID NOS: 14084
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1034
;   LENGTH: 686
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 2, 3, 8, 9, 12, 14, 16, 17, 20, 22, 23, 25, 27, 29, 31, 33,
;   LOCATION: 36, 37, 38, 43, 44, 45, 50, 53, 59, 63, 64, 65, 66, 83,
;   LOCATION: 85, 90, 93, 99, 106, 107, 192, 302, 303, 347, 368, 393,
;   LOCATION: 419, 470, 475, 496, 508, 543, 567, 571, 575, 580, 582, 587
;   OTHER INFORMATION: n = A,T,C or G
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 632, 639, 657, 665, 673
;   OTHER INFORMATION: n = A,T,C or G
US-10-198-846-1034

  Query Match             51.8%;  Score 20.2;  DB 5;  Length 686;
  Best Local Similarity   75.8%;  Pred. No. 1.3e+02;
  Matches   25;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          2 TAGCGAGCGTCTAATTCTAGAACACTCACACGC 34
              | | |||| |||| ||||||||||   | ||||
Db        220 TTGGGAGCCTCTATTTCTAGAACAGGAAAACGC 252


RESULT 9
US-10-779-543-23653/c
; Sequence 23653, Application US/10779543
; Publication No. US20050227917A1
; GENERAL INFORMATION:
;  APPLICANT: Williams et al
;  TITLE OF INVENTION: GENE PRODUCTS DIFFERENTIALLY EXPRESSED
;  TITLE OF INVENTION:  IN CANCEROUS CELLS AND THEIR METHODS OF USE II
;  FILE REFERENCE: 2300-21302
;  CURRENT APPLICATION NUMBER: US/10/779,543
;  CURRENT FILING DATE:  2004-02-12
;  PRIOR APPLICATION NUMBER: 10/076,555
;  PRIOR FILING DATE: 2002-02-15
;  PRIOR APPLICATION NUMBER: 09/217,471
;  PRIOR FILING DATE: 1998-12-21
;  PRIOR APPLICATION NUMBER: 60/068,755
;  PRIOR FILING DATE: 1997-12-23
;  PRIOR APPLICATION NUMBER: 60/080,664
;  PRIOR FILING DATE: 1998-04-03
;  PRIOR APPLICATION NUMBER: 60/105,234
;  PRIOR FILING DATE: 1998-10-21
;  PRIOR APPLICATION NUMBER: 09/297,648
;  PRIOR FILING DATE: 2000-04-10
;  PRIOR APPLICATION NUMBER: PCT/US99/01619
;  PRIOR FILING DATE: 1999-01-28
;  PRIOR APPLICATION NUMBER: 60/072,910
;  PRIOR FILING DATE: 1998-01-28
;  PRIOR APPLICATION NUMBER: 60/075,954
;  PRIOR FILING DATE: 1998-02-24
;  PRIOR APPLICATION NUMBER: 60/080,114
;  PRIOR FILING DATE: 1998-03-31
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 23767
```

```
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 23653
;   LENGTH: 773
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 640
;   OTHER INFORMATION: n = A,T,C or G
US-10-779-543-23653

  Query Match             51.8%;  Score 20.2;  DB 9;  Length 773;
  Best Local Similarity   75.8%;  Pred. No. 1.4e+02;
  Matches   25;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          2 TAGCGAGCGTCTAATTCTAGAACACTCACACGC 34
              | | |||| |||| ||||||||||   | ||||
Db        151 TTGGGAGCCTCTATTTCTAGAACAGGAAAACGC 119


RESULT 10
US-10-098-841-221/c
; Sequence 221, Application US/10098841
; Publication No. US20020197679A1
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Xu, Chongjun
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Ma, Yunqing
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Chen, Rui-hong
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Wang, Zhiwei
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Qian, Xiaohong B.
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: No. US20020197679A1el Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 784CIP2
;  CURRENT APPLICATION NUMBER: US/10/098,841
;  CURRENT FILING DATE:  2002-03-13
;  PRIOR APPLICATION NUMBER: 09/598,042
;  PRIOR FILING DATE: 2000-06-20
;  PRIOR APPLICATION NUMBER: 09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 331
;  SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 221
;   LENGTH: 2844
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (355)..(2223)
US-10-098-841-221

  Query Match             51.8%;  Score 20.2;  DB 5;  Length 2844;
  Best Local Similarity   75.8%;  Pred. No. 1.8e+02;
  Matches   25;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          2 TAGCGAGCGTCTAATTCTAGAACACTCACACGC 34
              | | |||| |||| ||||||||||   | ||||
Db       2476 TTGGGAGCCTCTATTTCTAGAACAGGAAAACGC 2444


RESULT 11
US-10-788-792-120/c
; Sequence 120, Application US/10788792
; Publication No. US20040191819A1
; GENERAL INFORMATION:
;  APPLICANT: Bayer Pharmaceuticals Corporation
;  APPLICANT:  Eveleigh, Deepa
;  APPLICANT:  Bigwood, Douglas
;  TITLE OF INVENTION: EXPRESSION PROFILES FOR BREAST CANCER AND METHODS OF USE
;  FILE REFERENCE: 5152
;  CURRENT APPLICATION NUMBER: US/10/788,792
;  CURRENT FILING DATE:  2004-02-27
;  PRIOR APPLICATION NUMBER: US 60/450,655
;  PRIOR FILING DATE: 2003-02-28
;  NUMBER OF SEQ ID NOS: 254
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 120
;   LENGTH: 2871
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-788-792-120

  Query Match             51.8%;  Score 20.2;  DB 8;  Length 2871;
  Best Local Similarity   75.8%;  Pred. No. 1.8e+02;
  Matches   25;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          2 TAGCGAGCGTCTAATTCTAGAACACTCACACGC 34
              | | |||| |||| ||||||||||   | ||||
Db       2112 TTGGGAGCCTCTATTTCTAGAACAGGAAAACGC 2080


RESULT 12
US-10-087-192-635/c
; Sequence 635, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
;  APPLICANT: Morris, David W.
;  APPLICANT:  Engelhard, Eric K.
;  TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
;  TITLE OF INVENTION:  CANCER
;  FILE REFERENCE: 529452000122
;  CURRENT APPLICATION NUMBER: US/10/087,192
;  CURRENT FILING DATE:  2002-03-01
;  PRIOR APPLICATION NUMBER: US 09/747,377
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: US 09/798,586
;  PRIOR FILING DATE: 2001-03-02
;  NUMBER OF SEQ ID NOS: 2059
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 635
;   LENGTH: 2881
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-087-192-635

  Query Match             51.8%;  Score 20.2;  DB 5;  Length 2881;
  Best Local Similarity   75.8%;  Pred. No. 1.8e+02;
  Matches   25;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          2 TAGCGAGCGTCTAATTCTAGAACACTCACACGC 34
              | | |||| |||| ||||||||||   | ||||
Db       2122 TTGGGAGCCTCTATTTCTAGAACAGGAAAACGC 2090


RESULT 13
US-10-058-270A-65/c
; Sequence 65, Application US/10058270A
; Publication No. US20040029114A1
; GENERAL INFORMATION:
;  APPLICANT: Mack, David H.
;  APPLICANT:  Gish, Kurt C.
;  APPLICANT:  Afar, Daniel
;  APPLICANT:  Eos Biotechnology, Inc.
;  TITLE OF INVENTION: Methods of Diagnosis of Breast Cancer, Compositions and
```

```
;  TITLE OF INVENTION:  Methods of Screening for Modulators of Breast Cancer
;  FILE REFERENCE: 018501-005210US
;  CURRENT APPLICATION NUMBER: US/10/058,270A
;  CURRENT FILING DATE:  2002-01-24
;  PRIOR APPLICATION NUMBER: US 60/263,965
;  PRIOR FILING DATE: 2001-01-24
;  PRIOR APPLICATION NUMBER: US 60/265,928
;  PRIOR FILING DATE: 2001-02-02
;  PRIOR APPLICATION NUMBER: US 09/829,472
;  PRIOR FILING DATE: 2001-04-09
;  PRIOR APPLICATION NUMBER: US 60/282,698
;  PRIOR FILING DATE: 2001-04-09
;  PRIOR APPLICATION NUMBER: US 60/288,590
;  PRIOR FILING DATE: 2001-05-04
;  PRIOR APPLICATION NUMBER: US 60/294,443
;  PRIOR FILING DATE: 2001-05-29
;  NUMBER OF SEQ ID NOS: 141
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 65
;   LENGTH: 2881
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-058-270A-65

  Query Match             51.8%;  Score 20.2;  DB 7;  Length 2881;
  Best Local Similarity   75.8%;  Pred. No. 1.8e+02;
  Matches   25;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          2 TAGCGAGCGTCTAATTCTAGAACACTCACACGC 34
              | | |||| |||| ||||||||||   | ||||
Db       2122 TTGGGAGCCTCTATTTCTAGAACAGGAAAACGC 2090


RESULT 14
US-10-198-846-10349/c
; Sequence 10349, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
;  APPLICANT: Lillie, James
;  APPLICANT:  Xu, Yongyao
;  APPLICANT:  Wang, Youzhen
;  APPLICANT:  Steinmann, Kathleen
;  TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
;  TITLE OF INVENTION:  FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
;  TITLE OF INVENTION:  THERAPY OF BREAST CANCER
;  FILE REFERENCE: MRI-049
;  CURRENT APPLICATION NUMBER: US/10/198,846
;  CURRENT FILING DATE:  2002-07-18
;  PRIOR APPLICATION NUMBER: 60/306,220
;  PRIOR FILING DATE: 2001-07-18
;  NUMBER OF SEQ ID NOS: 14084
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10349
;   LENGTH: 3230
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 3212, 3213, 3214, 3215, 3216, 3217, 3218, 3219, 3220, 3221,
;   LOCATION: 3222, 3223, 3224, 3225, 3226, 3227, 3228, 3229, 3230
;   OTHER INFORMATION: n = A,T,C or G
US-10-198-846-10349

  Query Match             51.8%;  Score 20.2;  DB 5;  Length 3230;
  Best Local Similarity   75.8%;  Pred. No. 1.9e+02;
  Matches   25;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          2 TAGCGAGCGTCTAATTCTAGAACACTCACACGC 34
              | | |||| |||| ||||||||||   | ||||
Db       2471 TTGGGAGCCTCTATTTCTAGAACAGGAAAACGC 2439


RESULT 15
US-10-087-192-634/c
; Sequence 634, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
;  APPLICANT: Morris, David W.
;  APPLICANT:  Engelhard, Eric K.
;  TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
;  TITLE OF INVENTION:  CANCER
;  FILE REFERENCE: 529452000122
;  CURRENT APPLICATION NUMBER: US/10/087,192
;  CURRENT FILING DATE:  2002-03-01
;  PRIOR APPLICATION NUMBER: US 09/747,377
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: US 09/798,586
;  PRIOR FILING DATE: 2001-03-02
;  NUMBER OF SEQ ID NOS: 2059
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 634
;   LENGTH: 46404
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(46404)
;   OTHER INFORMATION: n = A,T,C or G
US-10-087-192-634

  Query Match             51.8%;  Score 20.2;  DB 5;  Length 46404;
  Best Local Similarity   75.8%;  Pred. No. 3.4e+02;
  Matches   25;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          2 TAGCGAGCGTCTAATTCTAGAACACTCACACGC 34
              | | |||| |||| ||||||||||   | ||||
Db      35645 TTGGGAGCCTCTATTTCTAGAACAGGAAAACGC 35613


Search completed: January 30, 2006, 05:23:22
Job time : 350 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: January 30, 2006, 03:26:26 ; Search time 392.5 Seconds
(without alignments)
82.538 Million cell updates/sec

Title: US-10-663-033-REVCOMP4-SEQ3
Perfect score: 39
Sequence: 1 atagcgagcgtctaattctagaacactcacacgcgtctt 39

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 0.1

Searched: 6059551 seqs, 415333918 residues

Total number of hits satisfying chosen parameters: 12119102

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA_New:*
1: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq1:*
8: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq2:*
10: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq3:*
11: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	20.2	51.8	3342	7	US-10-750-185-34911	Sequence 34911, A
	2	20.2	51.8	3342	7	US-10-750-623-34911	Sequence 34911, A
	3	19.6	50.3	385	8	US-11-128-061-2345	Sequence 2345, Ap
	4	19.6	50.3	385	8	US-11-128-061-5987	Sequence 5987, Ap
	5	19.6	50.3	385	8	US-11-128-049-2345	Sequence 2345, Ap
	6	19.6	50.3	385	8	US-11-128-049-5987	Sequence 5987, Ap
	7	19.6	50.3	142605	8	US-11-121-086-64	Sequence 64, Appl
	8	19.2	49.2	127917	7	US-10-775-169-82	Sequence 82, Appl
c	9	19	48.7	1419	7	US-10-750-185-46559	Sequence 46559, A
c	10	19	48.7	1419	7	US-10-750-623-46559	Sequence 46559, A
	11	18.8	48.2	201	7	US-10-995-561-64263	Sequence 64263, A
	12	18.8	48.2	201	7	US-10-995-561-64296	Sequence 64296, A
	13	18.8	48.2	56448	7	US-10-995-561-13369	Sequence 13369, A
	14	18.6	47.7	930	8	US-11-055-822-509	Sequence 509, App
c	15	18.6	47.7	169725	8	US-11-121-086-63	Sequence 63, Appl
c	16	18.4	47.2	1981	7	US-10-750-185-58250	Sequence 58250, A
c	17	18.4	47.2	1981	7	US-10-750-623-58250	Sequence 58250, A
c	18	18.2	46.7	1124	7	US-10-750-185-24993	Sequence 24993, A
c	19	18.2	46.7	1124	7	US-10-750-623-24993	Sequence 24993, A
	20	18.2	46.7	1836	7	US-10-750-185-57119	Sequence 57119, A
	21	18.2	46.7	1836	7	US-10-750-623-57119	Sequence 57119, A
	22	18.2	46.7	3581	7	US-10-750-185-35045	Sequence 35045, A
	23	18.2	46.7	3581	7	US-10-750-623-35045	Sequence 35045, A
c	24	18.2	46.7	4232	8	US-11-024-959-100	Sequence 100, App
	25	18.2	46.7	179777	8	US-11-121-086-106	Sequence 106, App
	26	18.2	46.7	197096	8	US-11-121-086-107	Sequence 107, App
c	27	18	46.2	634	7	US-10-667-295-32	Sequence 32, Appl
	28	18	46.2	1161	7	US-10-750-185-35851	Sequence 35851, A
	29	18	46.2	1161	7	US-10-750-623-35851	Sequence 35851, A
c	30	18	46.2	1198	7	US-10-750-185-27263	Sequence 27263, A
c	31	18	46.2	1198	7	US-10-750-623-27263	Sequence 27263, A
	32	18	46.2	1711	7	US-10-750-185-63215	Sequence 63215, A
	33	18	46.2	1711	7	US-10-750-623-63215	Sequence 63215, A
	34	18	46.2	3223	7	US-10-750-185-27501	Sequence 27501, A
	35	18	46.2	3223	7	US-10-750-623-27501	Sequence 27501, A
c	36	18	46.2	398287	7	US-10-995-561-13396	Sequence 13396, A
	37	17.8	45.6	201	7	US-10-995-561-30958	Sequence 30958, A
	38	17.8	45.6	201	7	US-10-995-561-33376	Sequence 33376, A
	39	17.8	45.6	201	7	US-10-995-561-33378	Sequence 33378, A
	40	17.8	45.6	201	7	US-10-995-561-33395	Sequence 33395, A
	41	17.8	45.6	201	7	US-10-995-561-37693	Sequence 37693, A
	42	17.8	45.6	201	7	US-10-995-561-37895	Sequence 37895, A
	43	17.8	45.6	201	7	US-10-995-561-37896	Sequence 37896, A
c	44	17.8	45.6	600	7	US-10-750-185-525	Sequence 525, App
c	45	17.8	45.6	600	7	US-10-750-623-525	Sequence 525, App

ALIGNMENTS

RESULT 1
US-10-750-185-34911
; Sequence 34911, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DeNISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MMI1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 34911
; LENGTH: 3342
; TYPE: DNA
; ORGANISM: Bovine 19866880569723
US-10-750-185-34911

Query Match 51.8%; Score 20.2; DB 7; Length 3342;
Best Local Similarity 75.8%; Pred. No. 9.7;
Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```
Qy          4 GCGAGCGTCTAATTCTAGAACACTCACACGCGT 36
              |||    |  |||||| ||||||||||| ||||
Db       2328 GCGTCTTTAGAATTCTGGAACACTCACAAGCGT 2360
```

RESULT 2
US-10-750-623-34911
; Sequence 34911, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DeNISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom

```
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-1
;  CURRENT APPLICATION NUMBER: US/10/750,623
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 34911
;   LENGTH: 3342
;   TYPE: DNA
;   ORGANISM: Bovine  19866880569723
US-10-750-623-34911

  Query Match             51.8%;  Score 20.2;  DB 7;  Length 3342;
  Best Local Similarity   75.8%;  Pred. No. 9.7;
  Matches   25;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          4 GCGAGCGTCTAATTCTAGAACACTCACACGCGT 36
              |||    |  |||||| |||||||||||| ||||
Db       2328 GCGTCTTTAGAATTCTGGAACACTCACAAGCGT 2360


RESULT 3
US-11-128-061-2345
; Sequence 2345, Application US/11128061
; Publication No. US20060003958A1
; GENERAL INFORMATION:
;  APPLICANT: Melville, Mark W.
;  APPLICANT:  Charlebois, Timothy S.
;  APPLICANT:  Mounts, William M.
;  APPLICANT:  Hann, Louane E.
;  APPLICANT:  Sinacore, Martin S.
;  APPLICANT:  Leonard, Mark W.
;  APPLICANT:  Brown, Eugene L.
;  APPLICANT:  Miller, Christopher P.
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES RELATED TO OLIGONUCLEOTIDE ARRAYS
;  TITLE OF INVENTION:  TO MONITOR GENE EXPRESSION
;  FILE REFERENCE: 01997.027701
;  CURRENT APPLICATION NUMBER: US/11/128,061
;  CURRENT FILING DATE:  2005-05-11
;  PRIOR APPLICATION NUMBER: US 60/570,425
;  PRIOR FILING DATE: 2004-05-11
;  NUMBER OF SEQ ID NOS: 7285
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2345
;   LENGTH: 385
;   TYPE: DNA
;   ORGANISM: Cricetulus griseus
US-11-128-061-2345

  Query Match             50.3%;  Score 19.6;  DB 8;  Length 385;
  Best Local Similarity   73.5%;  Pred. No. 12;
  Matches   25;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          6 GAGCGTCTAATTCTAGAACACTCACACGCGTCTT 39
              |||  | | ||||| | |||||||||| | ||| |
Db        105 GAGAATTTCATTCTGGGACACTCACATGAGTCCT 138


RESULT 4
US-11-128-061-5987
; Sequence 5987, Application US/11128061
; Publication No. US20060003958A1
; GENERAL INFORMATION:
;  APPLICANT: Melville, Mark W.
;  APPLICANT:  Charlebois, Timothy S.
;  APPLICANT:  Mounts, William M.
;  APPLICANT:  Hann, Louane E.
;  APPLICANT:  Sinacore, Martin S.
;  APPLICANT:  Leonard, Mark W.
;  APPLICANT:  Brown, Eugene L.
;  APPLICANT:  Miller, Christopher P.
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES RELATED TO OLIGONUCLEOTIDE ARRAYS
;  TITLE OF INVENTION:  TO MONITOR GENE EXPRESSION
;  FILE REFERENCE: 01997.027701
;  CURRENT APPLICATION NUMBER: US/11/128,061
;  CURRENT FILING DATE:  2005-05-11
;  PRIOR APPLICATION NUMBER: US 60/570,425
;  PRIOR FILING DATE: 2004-05-11
;  NUMBER OF SEQ ID NOS: 7285
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 5987
;   LENGTH: 385
;   TYPE: DNA
;   ORGANISM: Cricetulus griseus
US-11-128-061-5987

  Query Match             50.3%;  Score 19.6;  DB 8;  Length 385;
  Best Local Similarity   73.5%;  Pred. No. 12;
  Matches   25;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          6 GAGCGTCTAATTCTAGAACACTCACACGCGTCTT 39
              |||  | | ||||| | |||||||||| | ||| |
Db        105 GAGAATTTCATTCTGGGACACTCACATGAGTCCT 138


RESULT 5
US-11-128-049-2345
; Sequence 2345, Application US/11128049
; Publication No. US20060010513A1
; GENERAL INFORMATION:
;  APPLICANT: Melville, Mark W.
;  APPLICANT:  Charlebois, Timothy S.
;  APPLICANT:  Mounts, William M.
;  APPLICANT:  Hann, Louane E.
;  APPLICANT:  Sinacore, Martin S.
;  APPLICANT:  Leonard, Mark W.
;  APPLICANT:  Brown, Eugene L.
;  APPLICANT:  Miller, Christopher P.
;  TITLE OF INVENTION: OLIGONUCLEOTIDE ARRAYS TO MONITOR GENE EXPRESSION AND METHODS FOR
;  TITLE OF INVENTION:  MAKING AND USING SAME
;  FILE REFERENCE: 01997.027700
;  CURRENT APPLICATION NUMBER: US/11/128,049
;  CURRENT FILING DATE:  2005-05-11
;  PRIOR APPLICATION NUMBER: US 60/570,425
;  PRIOR FILING DATE: 2004-05-11
;  NUMBER OF SEQ ID NOS: 7285
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2345
;   LENGTH: 385
;   TYPE: DNA
;   ORGANISM: Cricetulus griseus
US-11-128-049-2345

  Query Match             50.3%;  Score 19.6;  DB 8;  Length 385;
  Best Local Similarity   73.5%;  Pred. No. 12;
  Matches   25;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          6 GAGCGTCTAATTCTAGAACACTCACACGCGTCTT 39
              |||  | | ||||| | |||||||||| | ||| |
Db        105 GAGAATTTCATTCTGGGACACTCACATGAGTCCT 138


RESULT 6
US-11-128-049-5987
; Sequence 5987, Application US/11128049
; Publication No. US20060010513A1
; GENERAL INFORMATION:
;  APPLICANT: Melville, Mark W.
;  APPLICANT:  Charlebois, Timothy S.
;  APPLICANT:  Mounts, William M.
;  APPLICANT:  Hann, Louane E.
;  APPLICANT:  Sinacore, Martin S.
```

```
;  APPLICANT:  Leonard, Mark W.
;  APPLICANT:  Brown, Eugene L.
;  APPLICANT:  Miller, Christopher P.
;  TITLE OF INVENTION: OLIGONUCLEOTIDE ARRAYS TO MONITOR GENE EXPRESSION AND METHODS FOR
;  TITLE OF INVENTION:  MAKING AND USING SAME
;  FILE REFERENCE: 01997.027700
;  CURRENT APPLICATION NUMBER: US/11/128,049
;  CURRENT FILING DATE:  2005-05-11
;  PRIOR APPLICATION NUMBER: US 60/570,425
;  PRIOR FILING DATE: 2004-05-11
;  NUMBER OF SEQ ID NOS: 7285
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 5987
;   LENGTH: 385
;   TYPE: DNA
;   ORGANISM: Cricetulus griseus
US-11-128-049-5987

  Query Match             50.3%;  Score 19.6;  DB 8;  Length 385;
  Best Local Similarity   73.5%;  Pred. No. 12;
  Matches   25;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          6 GAGCGTCTAATTCTAGAACACTCACACGCGTCTT 39
              |||  | | ||||| | |||||||||| | ||| |
Db        105 GAGAATTTCATTCTGGGACACTCACATGAGTCCT 138


RESULT 7
US-11-121-086-64
; Sequence 64, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
;  APPLICANT: POULSEN, TIM S.
;  APPLICANT:  NIELSEN, KIRSTEN V.
;  TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
;  FILE REFERENCE: 09138.6000-00000
;  CURRENT APPLICATION NUMBER: US/11/121,086
;  CURRENT FILING DATE:  2005-05-04
;  PRIOR APPLICATION NUMBER: 60/567,570
;  PRIOR FILING DATE: 2004-05-04
;  NUMBER OF SEQ ID NOS: 107
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 64
;   LENGTH: 142605
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-121-086-64

  Query Match             50.3%;  Score 19.6;  DB 8;  Length 142605;
  Best Local Similarity   84.6%;  Pred. No. 41;
  Matches   22;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          7 AGCGTCTAATTCTAGAACACTCACAC 32
              |||||||||||| |||||| ||  |||
Db     117976 AGCGTCTAATTTTAGAACTCTTGCAC 118001


RESULT 8
US-10-775-169-82
; Sequence 82, Application US/10775169
; Publication No. US20050287532A9
; GENERAL INFORMATION:
;  APPLICANT: Wyeth
;  APPLICANT:  Burczynski, Michael
;  APPLICANT:  Twine, Natalie
;  APPLICANT:  Dorner, Andrew
;  APPLICANT:  Trepicchio, William
;  TITLE OF INVENTION: Method for Monitoring Drug Activities In Vivo
;  FILE REFERENCE: AM101080 (031896-013000)
;  CURRENT APPLICATION NUMBER: US/10/775,169
;  CURRENT FILING DATE:  2004-02-11
;  NUMBER OF SEQ ID NOS: 5278
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 82
;   LENGTH: 127917
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-775-169-82

  Query Match             49.2%;  Score 19.2;  DB 7;  Length 127917;
  Best Local Similarity   87.5%;  Pred. No. 62;
  Matches   21;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         15 ATTCTAGAACACTCACACGCGTCT 38
              || |||||||||||||  ||||||
Db      40267 ATGCTAGAACACTCACCAGCGTCT 40290


RESULT 9
US-10-750-185-46559/c
; Sequence 46559, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-2
;  CURRENT APPLICATION NUMBER: US/10/750,185
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 46559
;   LENGTH: 1419
;   TYPE: DNA
;   ORGANISM: Bovine  19866881196932
US-10-750-185-46559

  Query Match             48.7%;  Score 19;  DB 7;  Length 1419;
  Best Local Similarity   71.4%;  Pred. No. 29;
  Matches   25;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          4 GCGAGCGTCTAATTCTAGAACACTCACACGCGTCT 38
              ||   | | |||||||| | ||| |||| || |||
Db        597 GCATTCATATAATTCTAAAGCACACACATGCATCT 563


RESULT 10
US-10-750-623-46559/c
; Sequence 46559, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-1
;  CURRENT APPLICATION NUMBER: US/10/750,623
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 46559
;   LENGTH: 1419
;   TYPE: DNA
```

```
;   ORGANISM: Bovine  19866881196932
US-10-750-623-46559

  Query Match             48.7%;  Score 19;  DB 7;  Length 1419;
  Best Local Similarity   71.4%;  Pred. No. 29;
  Matches   25;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          4 GCGAGCGTCTAATTCTAGAACACTCACACGCGTCT 38
              ||   | | |||||||| | ||| |||| || |||
Db        597 GCATTCATATAATTCTAAAGCACACACATGCATCT 563


RESULT 11
US-10-995-561-64263
; Sequence 64263, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 64263
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-995-561-64263

  Query Match             48.2%;  Score 18.8;  DB 7;  Length 201;
  Best Local Similarity   76.7%;  Pred. No. 24;
  Matches   23;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          6 GAGCGTCTAATTCTAGAACACTCACACGCG 35
              |||| ||   |||| |||||||| |||| |
Db         68 GAGCCTCAGGTTCTGGAACACTCCCACGGG 97


RESULT 12
US-10-995-561-64296
; Sequence 64296, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 64296
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-995-561-64296

  Query Match             48.2%;  Score 18.8;  DB 7;  Length 201;
  Best Local Similarity   76.7%;  Pred. No. 24;
  Matches   23;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          6 GAGCGTCTAATTCTAGAACACTCACACGCG 35
              |||| ||   |||| |||||||| |||| |
Db         35 GAGCCTCAGGTTCTGGAACACTCCCACGGG 64


RESULT 13
US-10-995-561-13369
; Sequence 13369, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13369
;   LENGTH: 56448
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(56448)
;   OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-995-561-13369

  Query Match             48.2%;  Score 18.8;  DB 7;  Length 56448;
  Best Local Similarity   76.7%;  Pred. No. 80;
  Matches   23;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          6 GAGCGTCTAATTCTAGAACACTCACACGCG 35
              |||| ||   |||| |||||||| |||| |
Db      22983 GAGCCTCAGGTTCTGGAACACTCCCACGGG 23012


RESULT 14
US-11-055-822-509
; Sequence 509, Application US/11055822
; Publication No. US20050260707A1
; GENERAL INFORMATION:
;  APPLICANT: Pompejus, Markus
;  APPLICANT:  Kroger, Burkhard
;  APPLICANT:  Schroder, Hartwig
;  APPLICANT:  Zelder, Oskar
;  APPLICANT:  Haberhauer, Gregor
;  TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING
;  TITLE OF INVENTION:  METABOLIC PATHWAY PROTEINS
;  FILE REFERENCE: BGI-121CPCN
;  CURRENT APPLICATION NUMBER: US/11/055,822
;  CURRENT FILING DATE:  2005-02-11
;  PRIOR APPLICATION NUMBER: 09/606,740
;  PRIOR FILING DATE: 2000-06-23
;  PRIOR APPLICATION NUMBER: 60/141,031
;  PRIOR FILING DATE: 1999-06-25
;  PRIOR APPLICATION NUMBER: 60/142,101
;  PRIOR FILING DATE: 1999-07-02
;  PRIOR APPLICATION NUMBER: 60/148,613
;  PRIOR FILING DATE: 1999-08-12
;  PRIOR APPLICATION NUMBER: 60/187,970
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: DE 19930476.9
;  PRIOR FILING DATE: 1999-07-01
;  PRIOR APPLICATION NUMBER: DE 19931415.2
;  PRIOR FILING DATE: 1999-07-08
;  PRIOR APPLICATION NUMBER: DE 19931418.7
;  PRIOR FILING DATE: 1999-07-08
;  PRIOR APPLICATION NUMBER: DE 19931419.5
;  PRIOR FILING DATE: 1999-07-08
;  PRIOR APPLICATION NUMBER: DE 19931420.9
;  PRIOR FILING DATE: 1999-07-08
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 1158
; SEQ ID NO 509
;   LENGTH: 930
;   TYPE: DNA
;   ORGANISM: Corynebacterium glutamicum
;   FEATURE:
;   NAME/KEY: CDS
```

```
;   LOCATION: (101)..(907)
;   OTHER INFORMATION: RXC02080
US-11-055-822-509

  Query Match             47.7%;  Score 18.6;  DB 8;  Length 930;
  Best Local Similarity   84.0%;  Pred. No. 41;
  Matches   21;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          3 AGCGAGCGTCTAATTCTAGAACACT 27
              ||||||| |||||||||||  |  ||
Db         12 AGCGAGCTTCTAATTCTAGCAAGCT 36


RESULT 15
US-11-121-086-63/c
; Sequence 63, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
;  APPLICANT: POULSEN, TIM S.
;  APPLICANT:  NIELSEN, KIRSTEN V.
;  TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
;  FILE REFERENCE: 09138.6000-00000
;  CURRENT APPLICATION NUMBER: US/11/121,086
;  CURRENT FILING DATE:  2005-05-04
;  PRIOR APPLICATION NUMBER: 60/567,570
;  PRIOR FILING DATE: 2004-05-04
;  NUMBER OF SEQ ID NOS: 107
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 63
;   LENGTH: 169725
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-121-086-63

  Query Match             47.7%;  Score 18.6;  DB 8;  Length 169725;
  Best Local Similarity   72.7%;  Pred. No. 1.2e+02;
  Matches   24;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          7 AGCGTCTAATTCTAGAACACTCACACGCGTCTT 39
              || ||||  ||| |||| |||||| | | | ||
Db      89829 AGGGTCTTTTTCAAGAAGACTCACCCTCTTATT 89797


Search completed: January 30, 2006, 05:36:33
Job time : 393.5 secs
```